

**O'Bry n, Barbara**

---

**From:** Navarro, Albert  
**Sent:** Monday, February 05, 2001 11:07 AM  
**To:** O'Bryen, Barbara  
**Subject:** 09/001,737

Mark Navarro  
1645  
306-3225  
8A15

Hello Barb,

Could I trouble you for a search and an interference search of:

SEQ ID NO: 7

Thanks

Mark

POINT OF CONTACT:  
BARB O'BRYEN  
TECH. INFORMATION SPECIALIST  
STIC CM1 12C14 308-4291

*BOB*  
*2-16-01*

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:51:12 ; Search time 1971.82 seconds  
(without alignments)  
4311.034 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661  
Sequence: 1 GAATTCGGCTTCATATGCA.....TGGCGGATMACCGAATTC 1661

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_cm:\*  
4: gb\_ov:\*  
5: gb\_pb:\*  
6: gb\_pl1:\*  
7: gb\_pl2:\*  
8: gb\_pr1:\*  
9: gb\_pr2:\*  
10: gb\_pr3:\*  
11: gb\_ro:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: em\_fun:\*  
15: em\_hum1:\*  
16: em\_hum2:\*  
17: em\_in:\*  
18: em\_cm:\*  
19: em\_or:\*  
20: em\_ov:\*  
21: em\_pat:\*  
22: em\_ph:\*  
23: em\_pl:\*  
24: em\_ro:\*  
25: em\_sts:\*  
26: em\_sy:\*  
27: em\_un:\*  
28: em\_vl:\*  
29: gb\_ba3:\*  
30: gb\_in1:\*  
31: gb\_in2:\*  
32: gb\_in3:\*  
33: gb\_pl3:\*  
34: gb\_pr4:\*  
35: em\_da1:\*  
36: em\_da2:\*  
37: em\_hlg1:\*  
38: em\_hlg2:\*  
39: em\_hlg3:\*  
40: em\_hlg4:\*  
41: em\_hlg5:\*  
42: em\_hlg6:\*  
43: em\_hlg7:\*

44: em\_hlg8:\*  
45: em\_hlg9:\*  
46: em\_hlg10:\*  
47: em\_hum3:\*  
48: em\_hum4:\*  
49: em\_hum5:\*  
50: em\_hum6:\*  
51: gb\_pr5:\*  
52: gb\_pr6:\*  
53: gb\_pr7:\*  
54: gb\_hlg1:\*  
55: gb\_hlg2:\*  
56: gb\_hlg3:\*  
57: gb\_hlg4:\*  
58: gb\_hlg5:\*  
59: gb\_hlg6:\*  
60: gb\_hlg7:\*  
61: gb\_hlg8:\*  
62: gb\_hlg9:\*  
63: gb\_hlg10:\*  
64: gb\_hlg11:\*  
65: gb\_hlg12:\*  
66: gb\_hlg13:\*  
67: gb\_hlg14:\*  
68: gb\_hlg15:\*  
69: gb\_hlg16:\*  
70: gb\_hlg17:\*  
71: gb\_hlg18:\*  
72: gb\_hlg19:\*  
73: gb\_hlg20:\*  
74: gb\_hlg21:\*  
75: gb\_hlg22:\*  
76: gb\_hlg23:\*  
77: gb\_hlg24:\*  
78: gb\_sts2:\*  
79: gb\_vil1:\*  
80: gb\_vil2:\*  
81: gb\_pat1:\*  
82: gb\_pat2:\*  
83: em\_hlg0:\*  
84: gb\_hlg24:\*  
85: gb\_pr8:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1413.4	85.1	2683	29	SPGROELGN
2	1030.6	62.0	2401	1	AF117741
3	948.6	57.1	2267	2	LI6ROESEL
4	840.4	50.6	2834	1	AB028452
5	838.8	50.5	2309	2	BACSGROESTL
6	807.4	48.6	2400	2	BACSGROESTL
7	807.4	48.6	2525	2	BACHSPA
8	807.4	48.6	17516	1	AB007637
9	807.4	48.6	213190	2	BSUB0004
10	799.2	48.1	2190	2	BACSGROEL
11	795.8	47.9	2047	1	AF259784
12	786.4	47.3	296950	2	AF001508
13	773	46.5	3128	1	AF010281
14	771	46.4	2753	1	AF214488
15	756.4	45.5	3885	1	AF269843
16	751.6	45.2	2187	29	SEU13618
17	739.8	44.5	2540	2	D55630
18	736.2	44.3	2675	1	AF031929
19	722.4	43.3	2712	29	STAHSP
20	704.2	42.4	4710	2	LI045241
21	688.8	41.5	3987	2	CPGROESTL
					X89236 S.pyogenes
					AF117741 Streptococ
					X71132 L.lactis gr
					AB028452 Bacillus
					LI0132 Bacillus st
					M01132 Bacillus su
					M04965 B.subtilis
					AB007637 Bacillus
					299107 Bacillus su
					D10972 Bacillus su
					AF259784 Bacillus
					AF001508 Bacillus
					AF010281 Lactobaci
					AF214488 Lactobaci
					AF269843 Staphyloc
					LI3618 Staphylococ
					D55630 Bacillus sp
					AF031929 Lactobaci
					D14711 Staphylococ
					U45241 Lawsonia in
					X62914 C.perfringe

22	680.2	41.0	1940	1	AF080547	AF080547 Clostridi
23	674.4	40.6	2204	1	CGROBSEL	Z68137 C. thermocel
24	673.8	40.6	1623	1	AF159449	AF159449 Clostridi
25	657	39.6	1644	2	BHU78514	U78514 Bartonella
26	656.4	39.5	2103	29	TBU56021	U56021 Thermomanaer
27	652.6	39.3	3493	2	CLOGROESIA	M4572 Clostridium
28	652.2	39.3	1644	2	ROU78515	U78515 Bartonella
29	646.4	38.9	1627	1	AF014829	AF014829 Bartonella
30	645.6	38.9	1894	2	BABOB63A	M98257 Bartonella
31	644.6	38.8	1811	2	LPNHtPB3	M31918 Legionella
32	643.4	38.7	1788	2	PAH243594	AJ243594 Pseudocalt
33	640.8	38.6	1635	2	BEGROELT	Z15160 B. bacilliflo
34	636.6	38.3	5907	2	D856528	D856528 Buchnera ap
35	636.6	38.3	346699	2	AF008210	AF008210 Buchnera
36	630.6	38.0	282183	1	CI11168X4	AI135077 Campyloba
37	629	37.9	2580	2	CYV13334	Y13334 Campylobact
38	615.8	37.1	2016	2	LPNHtPB	M1673 Legionella
39	615.8	37.1	2739	2	ASYSYSL	X61150 Acyrthosiph
40	615.8	37.1	11763	29	U32736	U32736 Haemophilus
41	615.8	37.1	347550	2	AP001118	AP001118 Buchnera
42	614.2	37.0	6224	1	AF003957	AF003957 Myzus per
43	611.4	36.8	2596	2	FTGRO	X88833 F. tularensi
44	609.4	36.7	1522	1	AF014833	AF014833 Bartonell
45	603	36.3	2360	2	PMU030165	U30165 Pasteurella

## ALIGNMENTS

AF0809547 Clostridi		AGIEGSAVIDIKLNKSPACTGFNNAITGEWMDIKTGIIIDPVKVTRSLAQNAAVASNLIL
Z68137 C. thermocel		TTEAVVANKPEPPAAPAPMAPMGDMGMGG"
AF159449 Clostrid	repeat_region	1509..1564
U78514 Bartonella		/rpl_type=INVERTED
U56021 Thermomast	BASE COUNT	808 a 490 c 593 g 792 t
M74572 Clostridium	ORIGIN	
U78515 Bartonella		
AF014829 Bartonell		
M98257 Bartonella	Query Match	85.1%; Score 1413.4; DB 29; Length 2683;
M31918 Legionella	Best Local Similarity	98.9%; Pred. No. 2.1e-257;
AJ243594 Pseudolat	Matches 1423; Conservative	0; Mismatches 16; Indels 0; Gaps 0;
Z15160 B.bacilliffo		
D85628 Buchnera ap		
AF008210 Buchnera	QY 201	GATCATTTGGAAAACATCGGGAGCAAAATTGGTGTCTGAAGTGGCTTCCTAAAAACAATAT 260
AL139077 Campyloba	Db 1	GATCATTTGGAAAACATCGGGAGCAAAATTGGTGTCTGAAGTGGCTTCCTAAAAACAATAT 60
Y13334 Campylolact		
M91673 Legionella	QY 261	ATTGCTGTGTATGGGACGACTACTGCACAACAGTTTGGACACAAGCCATTGCCATGAAGA 320
X61150 Acyrthosiph	Db 61	ATTGCTGTGTATGGGACGACTACTGCACAACAGTTTGGACACAAGCCATTGCCATGAAGA 120
U32736 Haemophilus		
AP001118 Buchnera		
AF003957 Myzus per	OY 321	CYRAAAATGTGACACAGCAGTGCTTAATCCAATTGTTATTCGTCGAGGCATTGGAACAGCA 380
X98853 F.tularensi		
AF014833 Bartonell	Db 121	CYRAAAATGTGACACAGCAGTGCTTAATCCAATTGTTATTCGTCGAGGCATTGGAACAGCA 180
Z30165 Pasteurella		

RESULT	1
SPGROELGN	LOCUS
SPGROELGN	2683 bp DNA
DEFINITION	S. pyogenes DNA for groEL gene.
ACCESSION	X89236
VERSION	X89236.1 GI:2462691
KEYWORDS	groEL gene; heat shock protein 60 (groEL) like protein.
SOURCE	Streptococcus pyogenes.
ORGANISM	Streptococcus pyogenes.
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
AUTHORS	Streptococcus.
TITLE	1 (bases 1 to 2683)
JOURNAL	Podbielski, A.
REFERENCE	Direct Submission
AUTHORS	Submitted (29-JUN-1995) A. Podbielski, Institut f Med
JOURNAL	Microbiologie, an der Rhein Westf. Techn. Hochschule, Aachen,
REFERENCE	Pohl, B., Podbielski, A. and Zarges, I.
AUTHORS	2 (bases 1 to 2683)
JOURNAL	Pohl, B., Podbielski, A. and Zarges, I.
REFERENCE	Unpublished
AUTHORS	Related sequences M81132, M84965.
JOURNAL	Location/Qualifiers
COMMENT	1. 2683
FEATURES	/organism="Streptococcus pyogenes"
SOURCE	/strain="serotype M49"
	/isolate="CS101"
	/db_xref="taxon:1314"
	1. 1446
gene	/gene="groEL"
	<1. 1446
CDS	/gene="groEL"
	/codon_start=1
	/transl_table=1
	/product="heat shock protein"
	/protein_id="CA61520.1"
	/db_xref="GI:2462692"
	/db_xref="SPTRMBL:O33733"
	/translation="DHPENNGAKLVSEVASKTNDIAGDGTATATVLTFOATYHEGLKVV
	TAGNPGIGRIQGITATATATAVEALAKAOPVSGKEALTAOYAAVSRSEKGEYTSMM
	ERKNGDVTITEESRGMETELFYVEGMOFDRGYSLQYMTYDNEKVAIDLNPFLIND
	KKYSINDIITPLEEYVETKTRNPILITADVDGEALPTVLNLRIGTINVAVKAPFG
	DRRKAMLEDIATLTGGVITLEDLDELKDDATMRLAGVAVIVKGAPETALKEMLRIED
	ALNARILRSQLETTSDPDREKLQERLAKLAGVAVIVKGAPETALKEMLRIED
	ALNARAAVEGIVAGGCTALITIEKVALALETGDDATGRNIVLRLALEEVRQIALN

Repeat_region		AGEGSVIDKLKNSPAGTGFENNAATGEWMDIKTGIIDPVKTVRSALONAA5V	
1509..1564		TTEAVVANKREPEAPAPAMPAGMDPEGMGF*	
/rpt_type=INVERTED			
BASE COUNT	808 a 490 c 593 g 792 t		
ORIGIN			
Query Match	85.1%; Score 1413.4; DB 29; Length 2683;		
Best Local Similarity	98.9%; Pred. No. 2,1e-257;		
Matches 1423; Conservative	0; Mismatches 16; Indels 0; Gaps		
OY	201	GATCATTTTGAACAATGGAGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAAACCAATGAT	260
Db	1	GATCATTTTGAACAATGGAGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAAACCAATGAT	60
OY	261	ATTGCTGTGTATGGGAGGACTACTGCAACAGTTTGTGACACAGCCATTGTTATGAGA	320
Db	61	ATTGCTGTGTATGGGAGGAGCTACTGCAACAGTTTGTGACACAGCCATTGTTATGAGA	120
OY	321	CTAAAAATGTGAAGCAGGAGTGCTAATCCAAATTTGGTATCCGTGAGGCAATTGAACAGA	380
Db	121	CTAAAAATGTGAAGCAGGAGTGCTAATCCAAATTTGGTATCCGTGAGGCAATTGAACAGA	180
OY	381	ACAGCAACAGCTGTTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAGAGAGCT	440
Db	181	ACAGCAACAGCTGTTTGAAGCCTTGAAAGCCCTTGCTCAACCTGTATCTGGCAAGAGAGCT	240
OY	441	ATTGCTCAGGTGCTGCGCATATCATCACGCTCTGAAAAAGTTGGAGATATATCTCAGAA	500
Db	241	ATTGCTCAGGTGCTGCGCATATCATCACGCTCTGAAAAAGTTGGAGATATATCTCAGAA	300
OY	501	GCTATGGAGCCTGTGGGCAACGATGGTGGTATACCATGGAGAAATCTGGAGTATGGAA	560
Db	301	GCTATGGAGCCTGTGGGCAACGATGGTGGTATATACCATGGAGAAATCTGGAGTATGGAA	360
OY	561	ACAGAACTTGAAGTGTGTGAAGGCATGCAATTTGACCCTGGTATCCTGTCTCAATAATCATG	620
Db	361	ACAGAACTTGAAGTGTGTGAAGGCATGCAATTTGACCCTGGTATCCTGTCTCAATAATCATG	420
OY	621	GTCACAGACAATGAAAAATGTTGCGAGACCTTGAAAAACCCATTATCTTAATCAACGAT	680
Db	421	GTCACAGACAATGAAAAATGTTGCGAGACCTTGAAAAACCCATTATCTTAATTTGAT	480
OY	681	AAAAAAGTCAAAATCCAAAGCAATTTTGGCCACTCTGAGSAAAGTCTTTAAACCAAC	740
Db	481	AAAAAAGTCAAAATCCAAAGCAATTTTGGCCACTCTGAGSAAAGTCTTTAAACCAAC	540
OY	741	CGTCCATTACTCATTTATGCAAGATGATGTGAGTGAAGAGCACTTCCAAACCTTGTCTTG	800
Db	541	CGTCCATTACTCATTTATGCAAGATGATGTGAGTGAAGAGCCCTTCCAAACCTTGTCTTG	600
OY	801	AACAAGATTGCTGCTACTTTCATGTGTGCTGTCTCAAAAGCCGACAGATTTGCTATCGT	860
Db	601	AACAAGATTGCTGCTACTTTCATGTGTGCTGTCTCAAAAGCCGACAGATTTGCTATCGT	660
OY	861	CGTAAAGCTATGCTTGAAGACATTTGCTATCTTACAGAGTGTGATGATGAT	920
Db	661	CGTAAAGCTATGCTTGAAGACATTTGCTATCTTACAGAGTGTGATGATGAT	720
OY	921	CTAGAGCTTGAATTTAAAAAGATGCTACAAATGACAGCCCTTGGACAGGCTCTAAAGATTACA	980
Db	721	CTAGAGCTTGAATTTAAAAAGATGCTACAAATGACAGCCCTTGGACAGGCTCTAAAGATTACA	780
OY	981	GTTGATTAAGATAGACACAGTAATTTGTTGAAGGTTTCAGAAAGTTCAAGAACATATTTGCTAAC	1040
Db	781	GTTGATTAAGATAGACACAGTAATTTGTTGAAGGTTTCAGAAAGTTCAAGAACATATTTGCTAAC	840
OY	1041	CGTATTCGACTGATTAATATGCGCAATTATGAAACCAACACTTTCGACTTTGACCGGTAAAAA	1100
Db	841	CGCATTCGACTGATTAATATGCGCAATTATGAAACCAACACTTTCGACTTTGACCGGTAAAAA	900
OY	1101	CTCAACAACAGTATTGGGAAATTAAGTGTGGTGTGTATGCTTATCAAGATAGAGCTCTCA	1160

Db	901	CTACAGAACGCTTTGGCCAAATTAGCTGCTGCTAGCTGTTATCAAAAGTAGAGCTCCA	960
QY	1161	ACAGAGACAGCTTTTAAAGAAATGAACCTTCGCAFTGAGAGTGCCTTAAATGCTACACGT	1220
Db	961	ACAGAGACAGCTTTTAAAGAAATGAACCTTCGCAFTGAGAGTGCCTTAAATGCTACACGT	1020
QY	1221	GCAGCCGTTGAAAGAAAGTATGCTGGTGGTGGTGGAAACAGCACTTATTACGGTATTGAA	1280
Db	1021	GCAGCCGTTGAAAGAAAGTATGCTGGTGGTGGTGGAAACAGCACTTATTACGGTATTGAA	1080
QY	1281	AAAGTAGCAGCTCTTGACCTTGAGGGGCGATGATCTACTGAGCTAACATTGCTCGT	1340
Db	1081	AAAGTAGCTGCTCTTGACCTTGAGGGGCGATGATCTACTGAGCTAACATTGCTCGT	1140
QY	1341	GCTCTAGAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGCCTCCGTAGTT	1400
Db	1141	GCTCTAGAGAGAGCCTGTACGTCAAATTGCTTTAAATGCTGGGTACGAAGCCTCCGTAGTT	1200
QY	1401	ATTGCAAGTGGAAAAACAGCCCTGCAGAGAACAGAGATTAAATGCTGCACAAAGTGAATGG	1460
Db	1201	ATTGCAAGTGGAAAAACAGCCCTGCAGAGAACAGAGATTAAATGCTGCACAAAGTGAATGG	1260
QY	1461	GTTGATTATGATTAACAACGAAATATATGACCCCTGCAAAGTAACACAGTACAGCGCTCAA	1520
Db	1261	GTTGATTATGATTAACAACGAAATATATGACCCCTGCAAAGTAACACAGTATACGCGCTTCAA	1320
QY	1521	AATGCAGCTTCTGTAGCTAGCTTATTTTGAACAACAGAACAGTTGTTGCTAAATAAACCT	1580
Db	1321	AATGCAGCTTCTGTAGCTAGCTTATTTTGAACAACAGAACAGTTGTTGCTAAATAAACCT	1380
QY	1581	GAACCAAGTACGCCAGCGCCACGCAATCCGACAGATATGATTCAGAAATGATGGGTGG	1639
Db	1381	GAACCAAGTACGCCAGCGCCACGCAATCCGACAGATATGATTCAGAAATGATGGGTGG	1439

RESULT	2
LOCUS	AF117741
DEFINITION	AF117741 2401 bp DNA BCPT 06-APR-1999.
ACCESSION	Streptococcus pneumoniae strain CP1200 cochaparonin GroES (groes) and chaparonin GroEL (groEL) genes, complete cds.
VERSION	AF117741
KEYWORDS	AF117741.1 GI:4566771
SOURCE	
ORGANISM	Streptococcus pneumoniae. Streptococcus pneumoniae Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE	1 (bases 1 to 2401) Kim,S.N., Lee,J.Y., Kim,S.W., Choi,I.H. and Rhee,D.K. groESL sequences in Streptococcus pneumoniae Unpublished
AUTHORS	2 (bases 1 to 2401) Kim,S.N., Kim,S.W., Choi,I.H. and Rhee,D.K. Direct Submission Submitted (04-JAN-1999) Pharmacy, Sungkyunkwan University, 300 Chunchun-dong, Suwon, Kyunggi-Do 440-746, Korea
TITLE	Location/Qualifiers
JOURNAL	1..2401
FEATURES	
source	1..2401 /organism="Streptococcus pneumoniae" /strain="CP1200" /db_xref="taxon:1313" 122..603 /gene="groES" 122..127 /gene="groES" 147..152 /gene="groES" 262..288 /note="controlling inverted repeat of chaparonin expression" /rpt_family="CIRCE" /rpt_type=inverted
gene	
-35_signal	
-10_signal	
repeat_region	

```

BRS      307. 311
         /gene="groes"
         319. .603
         /gene="groes"
         /codon_start=1
         /transl_table=1
         /product="cochaperonin Groes"
         /protein_id="AAD3454.1"
         /db_xref="GI:4566772"
         /translation="MLKPLGDRVLLKLEEKREQYVGGVLAGSQAERTKTAQVVAATGGG
         VRTINGDLVAPSVTGDRLVLEAHAGLDVKGDEKYLIIIGEANLALIEE"
gene      606. .2241
         /gene="groEL"
         606. .610
         /gene="groEL"
         619. .2241
         /gene="groEL"
         /codon_start=1
         /transl_table=1
         /product="chaperonin GroEL"
         /protein_id="AAD3455.1"
         /db_xref="GI:4566773"
         /translation="MSKEKFKSPDARSAMRGVDILADYKVTILGPKDRNVLEKSG
         SPLTINDGVTIAKETLEEDHFNENGAALVSEIASKTINDIGDGTTAATVLTQIVREG
         IKNTAGANPLGIRGILETAVAALVEALKNNALIPVANKKALISOVAASVSSSEVGEYI
         SEAKERQKGVDDVYIIIESRGMETLEVEEGMGPRGYLSQYMTDSEKMYADALEPYI
         LIIDPKKISINIDELPLLESILIQSNRPILLIADVDGALPLVILNKRIGFNVAVA
         PGFEDRRKAMEADIAILTGCTVIEDGLIDKDATIEALGOAARVYDKDSIVIEGA
         GNPALASHRAVAIVISOLETTTSEFDREKLOERLAKISGGVAIVKGAATFETLEKMKL
         RIEHALNRAVAEETIVAGGCTALANIVAEATIELTGDEATGRNVLALAEPPNO
         IAHNAGFEQSVIIDRLKNAELGIFSNATGEVAMIDOGIITDPVKVSRSLQNNASVA
         SLILITTEVVAANKPEVAPAPAMPDPSMMGGM"
terminator 743 a 459 c 568 g 631 t
          2274. .2302
ORIGIN

```

Query Match	Best Local Similarity	62.0%	Score 1030.6	DB 1	Length 2401
Matches 1252	Conservative	0	Mismatches 369	Indels	Gaps
0Y	14	TATGCAAAAGAAATCAATTTTCGACGATGCGGCTGCTCCATGGGCGCGAGTTGA	73		
	618	TATGCAAAAGAAATTAATTTTCATCGATGCCGCTGACGCCATGGTGTGGTGTGCA	677		
0Y	74	TATGTTACAGATACCGTCAAAAGTAACGCTTGTCCTAAAGGCGCAATGTTCTTTGA	133		
	678	TATCCTTGACAGACTGTAAAGTAACCTGGAGCAAAAGATGCAATGTGCTTTGA	737		
0Y	134	AAAGCTTTTGCTTCCCTTAATTACTAATGACGGGGTAACCATTTGCTAAAGATGCA	193		
	738	AAAGCATTTGGTGTCACCCCTTGATTCACCAATGACGGTGTGACCATGGCAAAAGATCA	797		
0Y	134	ATTAAAGTCAATTTTGAAGACATGGGAGCAAAATGCTGTCTGAATGGCTTTAAAC	253		
	788	ATTGGAAGACCATTTTGAAGAAATGAGGGCTTAAGTAACTATACAAAGTAAGCTTTAAAC	857		
0Y	254	CAATGATATTGCTGTGTGAGGACAGCAGTCTGCAACATTTTGAACACAGCATTTGTCA	313		
	858	CAATATATTCGAGGTGACGGGACTACAGCTGCAACAGCTTTGACCCAAAGTATTCGCG	917		
0Y	314	TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATTCAAATTGGTATCCGTCGAGGCATTGA	373		
	918	TGAAGGATTCAAAAACGTCACAGCAGGTGCAAAATCCAAATCGTATTCGTCGGGATTGA	977		
0Y	374	AACGCAACAGCAGCAGCTGTGTGAAGCCTTAAAGCCATTGCTCAACCTGATATGGCAA	433		
	978	AACAGCACTTCGCGAGGAGTGTGAAGCTTTTAAAAACACAGCCATCCCTGTGTGCCAATAA	1037		
0Y	434	GGAAGCTATTGCTGACGGTCTGCGAGTATCAATCAAGCGCTGAAAGAGTGGAGGATAT	493		
	1038	AGAGCTATTCCTCAAGTTGACGCGATATCTTCGTGTTTAAAGAGTGGTGAATCAT	1097		

QY	494	CTCAGAAGCTTGAGAGCGGTGTGGGACACAGTGGTGTGATTACATCGAAGAAATCTCGAGG	553
Db	1098	CTCTGAAAGCAATGAAAAAAGTTGGCCAAAGACGGTGTGATCACCATTGAGAGTCAACGTGG	1157
QY	554	TATGAAAAACAGAACTTGTAAGTGGTTGAAGGCATGCAATTTGACCGGTGTACTGTCTCA	613
Db	1158	TATGAAAAACAGAGCTTGAAGTCGTAGAGGAATGACAGTTTGAACCGGTGTACTCTTACA	1217
QY	614	ATCATGTGCTCACAGACATATAAAAATGGTGTGCAGACCTTGGAAAAACCTTTATCTTAT	673
Db	1218	GTCATGTGCTCATATAGCCAAAAAATGGTGGCTGACCTTGAATAATCCGTACATTTTGAT	1277
QY	674	CACGATATAAAAAAGTGCAAACATCCAAAGCATTTTGGCCATGACTTGAAGAAAGTTCTAA	733
Db	1278	TACAGACAGAAAAATTTCCATATTCACAGAAATCTTGGCAGCTTTGGAAAGCATTTCTCA	1337
QY	734	AACCAACCGTCCATTACTCATTTATTCAGATGATGTGATGGTGAAGACACTTCCAAACCT	793
Db	1338	AAGCAATGTCCACTCTGTATTTATGGGATGATGTGTGATGGCGAGGCTCTCCAAACCT	1397
QY	794	TGCTTTGACACAGATTCGTGTGCTTTCATATGTGGTGTCTTCCAAAGCCGACAGATTTGG	853
Db	1398	TGTTTTGAACAAGATTCGTGTGAACCTTCAACGATGAGCAGTCAAGGACACCTGGTTTTGG	1457
QY	854	TGATGTCGTGAAAAGCTATGGTTGACACATGTGTATCTTGACAGAGTGTACAGCATATAC	913
Db	1458	TGACCGTGCAGAAACCCATGGCTTGAATATATGCGCATCTTAAACGCGGAAACGTTATAC	1517
QY	914	AGAGGATCTAGACACTTGAATAATTAAGATGCTCAATATGACAGCGCTTGGACAGCGCTAA	973
Db	1518	AGAAGACCTTGGCTTGTAGTTGAAGATGACAAATATGAAGTCTTGGTCAAGACGCGAG	1577
QY	974	GATTACAGTTGATTAAGATATGCAACAGTAAATTTGTTGAAGTTTGAAGAGTTCAGAAGCTAT	1033
Db	1578	AGTACCGTGGACAAAGATATGACAGCGTTATTTGAGAGGTGCAGGAATAATCCGAAAGCAT	1637
QY	1034	TGCTAACCGTATTTGCACTGTTTAAATGCGAATTAGAAACAACAATCTTGACTTTGACCG	1093
Db	1638	TTCTCACCGGTGTGGCTTATCAAGTCTCAAAATCGAAATACAACTTCTGCAATTTTGACCG	1697
QY	1094	TGAAAAACCTCAAGAACGTTTGGCGAAATTTAGCTGGTGTAGCTGTTATCAAAATGTAG	1153
Db	1698	TGAAAATTTGCAAGAAACGCTTGGCCAAATTTGCAGTGGTGTAGGGGTATTAAAGTTGG	1757
QY	1154	AGCTCAACACAGACAGCTTTAAAGAAATGAACCTTGCATTTGAGATGCTCTAAATGC	1213
Db	1758	AGCCGCAACTGAAACCTGAGTTGAAGAAATGAACCTCCGATTTGAAGATGCCCTCAACGC	1817
QY	1214	TACACGTGCAGCCGTTGAAGAAAGCTTCTGCTGCTGGTGGTAACAGCACTTATTAGCGT	1273
Db	1818	TACTCTGTGCAGCTTTGAAGAAAGTATTTGTTCAGGTGGTGGAAACAGCTCTTGGCCAAATGT	1877
QY	1274	TATTTGAAAAAGTAGACAGCTCTTACGCTGTAGGGCGATGTGCTACTGCACTTAACATGT	1333
Db	1878	GATTCACAGCTGGAAGCTTACCTTGAATTTGACAGSAGATGAGCAACAGSAGCTAATATTGT	1937
QY	1334	GCTTCTGTCTTAGAAGAGCGCTGTACGTCAAAATGCTTTAAATGCTGGGTACGAAGCTC	1393
Db	1938	TCTCCGTGCTTTGAAGAAACCGCTGTGTCAAATTTGCTCACAATGACAGGATTTGGAAGATC	1997
QY	1394	CGTAGTTATTTGACAAAGTTGAAAAACAGCCCTGCAGAGAACGATTTAATGCTGCAACAG	1453
Db	1998	TATCTGTTATGATCGTTTGAAAAAATGCTGAGCTTGTATAGGATTTTAAACGACAGATCG	2057
QY	1454	TGAGTGGGTTGATATGATTTAAAAACAGGATCTATTAACCTGTGCAAGATGACAGCATGAGC	1513
Db	2058	CGAGTGGGTTTAACATGATTTGATCAAGGTATCTTTATATCCAGTTAAAGTGAAGTCTTGAC	2117
QY	1514	GCTTCAAAATGACAGCTTCTTAGCTAGCTATGCTTATTTTGTGCAACAGAAAGATTTGTTCTAA	1573
Db	2118	CTTACAAAATATGACAGCATGTGTAGCCAGCTTGTATTTTGAACAACAGAGCAGTGTAGCCAA	2177
QY	1574	TAAACCTGAAACAGCTTACGCCAGCGCCACCAATGCCAGAGGTATGATTCAGGAATGAT	1633

Db	2178	TAAACCAAGAACCGATGACCCCACTCCAGCAATGATCCCAACCAATGATGGCGGGATAT	2237
QY	1634	G	1634
Db	2238	G	2238

RESULT	3
LIGROESSEL	LIGROESSEL
LOCUS	2267 bp
DEFINITION	L.lactis groes and groel genes.
ACCESSION	X71132
VERSION	X71132.1 GI:287869
KEYWORDS	groel gene; groes gene; heat shock protein.
SOURCE	Lactococcus lactis.
ORGANISM	Lactococcus lactis

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 2267)  
Kim,S.G. and Batt,C.A.  
Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis*. . .  
Gene. In press  
2 (bases 1 to 2267)  
Batt,C.A.  
Direct Submission  
Submitted (04-MAR-1993) C.A. Batt, Cornell University, 107 Stocking  
Hall, Ithaca, NY 14853, USA

```

source      1..2267
             /organism="Lactococcus lactis"
             /strain="ssp_lactis"
             /isolate="CC9"
             /db_xref="taxon:1358"
misc_feature 194..220
             /note="heat shock motif"
gene         246..530
             /gene="groES"
CDS          246..530

```

```

/codon_start=1
/transl_table=11
/protein_id="CA50445.1"
/db_xref="GI:287870"
/translation="MLKPLERVLRYKKEEEKSMGCIVLTSASQEKPOFEEVAVG
GRNNHGTITSLPIKVGVDYVLFKFACTYKMOSEFLIKDLSLLAIVE"
618..2246
/gene="groEL"
618..2246
CDS

```

```

/codon_start=1
/transl_table=11
/protein_id="CMA50446.1"
/db_xref="GI:287871"
/db_xref="SWISS-Prot:P37282"
/translation="MSKDIKFSSDAPRTAMRGIDILADTVYKTLTGPGRNVLKESYGV
SPILTDNGVITIAKEIELEFHPENNGALITSEVASKTDIDCGNTFTATVLOALVREEL
LKVATGAPNVGIRGIELAETPAVAIKEMALIVHRKSALAQVATVSSSEKCEGETI
SDMERGSDGVATIEESKGMQETLDEVGQFGRGLISQTMVNTKEMVALLNPVTA
LIDKATISNIQETILPLEQILIKTRPLLIADVDGALPLTVLNTKGVNVALVPA

```

BASE COUNT	763	a	357	c	487	g	660	t
ORIGIN	GALLILEAVANNPEENAFMPDD-SMGAGMT							
Query Match	57.1%	Score	948.6;	DB	2;	Length	2267;	
Best Local Similarity	74.9%	Pred. No.	1.3e-169;					
Matches 1188; Conservative	0;	Mismatches	359;	Indels	0;	Gaps	0;	

QY 11 TCATATGGCAAGAAATCAAATTTTCAGAGATGCGGTGTCCTCATGGTGGCGGAGT 70  
 614 TAAATATGCAAAAGATATTAATTTTCAAGATGATAGAACAGCGATGATGGCTGGAGT 673  
 QY 71 TGAATGTGTAGCAGATACCGTCAAGATACCGTTGGTCTTAAGAGCGCAATGTTTCT 130  
 674 TGAATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733  
 QY 131 TGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 190  
 734 TGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 793  
 QY 191 CGAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
 794 TGAATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853  
 QY 251 AACCAATGATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 310  
 854 AACCAATGATATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 913  
 QY 311 TCATGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGAT 370  
 914 TCGTGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGAT 973  
 QY 371 TGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 430  
 974 TGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1033  
 QY 431 CAAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 490  
 1034 TAAATATGCAAAAGATATTAATTTTCAAGATGATGATGATGATGATGATGATGATGAT 1093  
 QY 491 TATCTGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGAT 550  
 1094 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1153  
 QY 551 AGTATGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 610  
 1154 AGGATGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1213  
 QY 611 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670  
 1214 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273  
 QY 671 AATCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
 1274 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1333  
 QY 731 TAAAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 790  
 1334 GAAAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1393  
 QY 791 CTTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 850  
 1394 GCTTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1453  
 QY 851 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910  
 1454 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1513  
 QY 911 TACAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 970  
 1514 TACAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1573  
 QY 971 TAAAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1030  
 1574 TAAAGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1633  
 QY 1031 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090  
 1634 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1693

QY 1091 CCGTGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1150  
 1694 TCGTGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1753  
 QY 1151 AGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1210  
 1754 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1813  
 QY 1211 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270  
 1814 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1873  
 QY 1271 GCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330  
 1874 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1933  
 QY 1331 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390  
 1934 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1993  
 QY 1391 CCGTGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1450  
 1994 TCGTGAAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 2053  
 QY 1451 AGTATGAGAGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGAT 1510  
 2054 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2113  
 QY 1511 AGGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1570  
 2114 AGCTTTTGTCTCCCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2173  
 QY 1571 TAAATATGCAAAAGATATTAATTTTCAAGATGATGATGATGATGATGATGATGATGAT 1630  
 2174 TAAATATGCAAAAGATATTAATTTTCAAGATGATGATGATGATGATGATGATGATGATGAT 2200

RESULT 4  
 AB028452 2834 bp DNA BCT 07-DEC-1999  
 LOCUS  
 DEFINITION Bacillus sp. MS gene for Cpn10 and Cpn60, complete cds.  
 ACCESSION AB028452  
 VERSION AB028452.1 GI:6526959  
 KEYWORDS Cpn60; Cpn10.  
 SOURCE  
 ORGANISM Bacillus sp. MS (strain:MS) DNA.  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/staphylococcus group; Bacillus.  
 REFERENCE  
 AUTHORS Kondo, A. and Yoshida, M.  
 TITLE Cpn60/10 from Bacillus strain MS  
 JOURNAL Published only in Database (1999) In press  
 REFERENCE  
 AUTHORS Kondo, A. and Yoshida, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JUN-1999) to the DDBJ/EMBL/Genbank databases. Akihiko Kondo, Kobe University, Department of Chemical Science and Engineering, 1-1 Rokkodai-cho Nada, Kobe, Hyogo 657-8501, Japan (E-mail: kondo@chem.kobe-u.ac.jp, URL: http://www.kobe-u.ac.jp/, Tel: 078-803-6196, Fax: 078-803-6206)  
 LOCATION/Organism  
 FEATURES  
 source  
 1. 2834  
 /organism="Bacillus sp. MS"  
 /strain="MS"  
 /db\_xref="taxon:96470"  
 534..818  
 /gene="Cpn10"  
 534..818  
 /gene="Cpn10"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Cpn10"  
 /protein\_id="BAA8109.1"

```
/db_xref="GI:6526960"
/translation="MLKPIGDRVIVIEETEEKTASGIVLPDTAKEKPOEGRVAVNGK
GRVDSGEVRAVEVEVDRIIFSKEYGKDYDKEYLILRESDLIAVIG"
901..2520
/gene="Cpn60"
901..2520
/gene="Cpn60"
/codon_start=1
/transl_table=11
/product="Cpn60"
/protein_id="BAA89110.1"
/db_xref="GI:6526961"
/translation="MAKEIKFSEBARMLRGYDKLADAVKYTLGFKRNVLEKREG
SPILINDGVYIAKEIELEDPEFENMGAKLVAEASKINDVAGDGTATVLAOAMIREG
LNKVNAGAMPGRKIGIEKAVAVAEIKAIISPIKESISIOVAISADEVEQOLI
AAEMERVDGVYTLSESKFTELDVVEGMOFDRGVSPYMTTDEKMEAVLENNYI
LITDKISNIDILPIEQVOOGKPLLIADIEBEALATLVNKLKGFPAVAVKA
PGGDRKAMIEDIALITGSEVISELSEIKSATIASGRASKVYVTEENTIVGA
GDSERIKARINIROLEETSEFDEKLOERLAKLGVRYIKVGAATELEKRL
RIEDALNSTRVAVEGIVAGGTALNNVKNVALEDEDAIGVIVRALEEPRQ
IAONAGLEGSIVIERLTERKPGIGENAAIGEVNDMEAGIVDPTRKVSALONAAVSA
AMLTTEEAVVADKPEENKCNAGMPDMGMM"
BASE COUNT 954 a 525 c 668 g 687 t
ORIGIN
Query Match 50.6%; Score 840.4; DB 1; Length 2834;
Best Local Similarity 71.0%; Pred. No. 3.4e-149;
Matches 1114; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
QY 14 TATGGCAAAAGAAATCAATTTTCAGCAGATGCGCGTGCCTAGTGGCGGAGTTGA 73
DB 900 TATGGCAAAAGAAATTAATTAATTCAGCAGAGAGCTGCGTGCATGTACGTGGTGGA 959
QY 74 TATGTTAGCAGATACCGTCAAGTGGTGGCTTAAGGGCGCATGTGTTCTTGA 133
DB 960 CAATTTAGCGATGCGATTAATAATTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1019
QY 134 AAAAGCTTTTGGTTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 193
DB 1020 GAAATAATTCGTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1079
QY 194 ATTGAAGATCATTTTGAAGACATGGGAGCAAAATTTGTTGAGTGGCTCTTAAC 233
DB 1080 ATTGAAGATCATTTTGAAGACATGGGAGCAAAATTTGTTGAGTGGCTCTTAAC 1139
QY 254 CAATGATTTGCTGATGATGAGGAGCACTACTGCAACAGTTTGAACACCAATTTGTTCA 313
DB 1140 AAACGATGTTGCTGATGATGAGGAGCACTACTGCAACAGTTTGAACACCAATTTGTTCA 1199
QY 314 TGAAGACTAAAAATGTGACAGCAGCGCTGAATTCATTTGTTATCCGTCGAGCAATTGA 373
DB 1200 CGAAGACTTGAAGAAACGTTACAGCTGCGCAAAACCAATGGCATTCGTTAAAGATTTGA 1259
QY 374 AAAGCAACGCAACAGCTGTTGAAGCTTGAAGCCTATGCTCAACCTTATCTGGCAA 433
DB 1260 AAAAGCGGTTGCTGATGATGAGGAGCACTACTGCAACAGTTTGAACACCAATTTGTTCA 1319
QY 434 GGAAGCTATTGCTCAGTGCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 493
DB 1320 AGAATGATTTGCTCAGTGCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1379
QY 494 CTCAGAACTATGAGCGCTGTGGCAACGATGTTGATTAACATCGAAGATTCGAGG 553
DB 1380 CCCAGAACTATGAGCGCTGTGGCAACGATGTTGATTAACATCGAAGATTCGAGG 1439
QY 554 TATGAAGACGAGCTTGAAGCTGTTGAAGCCTATGATGATGATGATGATGATGATGAT 613
DB 1440 CTTCACAGCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
QY 614 ATACATGTCACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673
DB 1500 ATACATGTCACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559
```

```
QY 674 CACGATMAAAAGTGTCAAAACATCCAAAGACATTTTCCACTACTTGAAGAGTTCTTAA 733
DB 1560 CACAGACAAAAAATCTCTAACAATCCAAAGACATCTTCCATATCTTACAAAGTTGTTCA 1619
QY 734 AACCAACCGTCCATTTACTCTTATTTGATGATGATGATGATGATGATGATGATGATGAT 793
DB 1620 AACCAACCGTCCATTTACTCTTATTTGATGATGATGATGATGATGATGATGATGATGAT 1679
QY 794 TGCTTGAACAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853
DB 1680 AGTAGTAAACAACTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
QY 854 TGATGCTGTAAGCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 913
DB 1740 TGATGCTGTAAGCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
QY 914 AGAGGATCTGAGCACTGATTAATTAAGATGCTACATGACAGCCCTTGAGACGCTGTAA 973
DB 1800 CGAAGAGCTGAGCACTGATTAATTAAGATGCTACATGACAGCCCTTGAGACGCTGTAA 1859
QY 974 GATTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033
DB 1860 AGTAGTAAACAACTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919
QY 1034 TGCTAACGTTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
DB 1920 TAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1979
QY 1094 TGAATAAATCAAGAAAGCTTTGCGGAAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
DB 1980 TGAATAAATCAAGAAAGCTTTGCGGAAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039
QY 1154 AGCTCCACAGAGACAGCTTTAAAGAAATGAACCTCCGATGAGATGCTCTTAATGC 1213
DB 2040 TGCAGCTACAGAAACAGATTTGAAGAACCAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2099
QY 1214 TACAGCTGACGCGTGTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
DB 2100 TACTCGCGCGTGTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2159
QY 1274 TATTGAAAAAGTGAAGCACTTTGAGCTGAGGCGATGATGCTACATGAGCTTAATGCT 1333
DB 2160 ATACAAATTAAGTGTGCTGCTGATGAGCAAGGCGATGAGCAAGTGTGTAATGCT 2219
QY 1334 GCTTCGCTGCTGATGAGAGCGCTGATGCTCAATTTGCTTTAAATGCTGGTACAGAGCTC 1393
DB 2220 TCTTCGCGCATTTGAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2279
QY 1394 GCTGATTTGATGAGTGAAGAAACAGCCCTGAGAGCAAGATTTAATGCTGCAACAG 1453
DB 2280 TGTCAATTTGATGAGTGAAGAAACAGCCCTGAGAGCAAGATTTAATGCTGCAACAG 2339
QY 1454 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1513
DB 2340 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2399
QY 1514 GCTTCAAAATGACGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1573
DB 2400 ACTTCAAAAGCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2459
QY 1574 TAAACCTGAA 1583
DB 2460 CAAACCAAGAA 2469
RESULT 5
BACSGROESL 2309 bp DNA BCT 22-FEB-2000
LOCUS Bacillus stearothermophilus groES (groES) and GroEL (groEL) genes,
DEFINITION complete cds.
ACCESSION L10132.1
VERSION L10132.1 GI:289298
```



KEYWORDS      chaperonin; groEL gene; groES gene; groESL operon.  
SOURCE        Bacillus stearothermophilus.  
ORGANISM      Bacillus stearothermophilus  
              Bacteria; Firmicutes; Bacillus/Clostridium group;  
              Bacillus/staphylococcus group; Bacillus.  
REFERENCE     1 (bases 1 to 2309)  
AUTHORS       Schoen, U. and Schumann, W.  
TITLE          Molecular cloning, sequencing and transcriptional analysis of the  
              groESL operon from Bacillus stearothermophilus  
JOURNAL        J. Bacteriol. 175, 2465-2469 (1993)  
MEDLINE        93224474  
FEATURES  
    source        location/Qualifiers  
                  1..2309  
                  /organism="Bacillus stearothermophilus"  
                  /strain="NUB36"  
                  /db\_xref="taxon:1422"  
                  249..533  
                  /gene="groES"  
                  249..533  
                  /gene="groES"  
                  /codon\_start=1  
                  /transl\_table=1  
                  /product="groES"  
                  /protein\_id="AA022751.2"  
                  /db\_xref="GI:7019626"  
                  /translacion="MLKPLGDRVIVIEIETEETKASIVLPDEKEKPOEGRVAVGK  
                  GRVDSGERVAPEVEVDRIIFSKYAGTEVKYDGEKVLILRESDIILAVIG"  
                  606..2225  
                  /gene="groEL"  
                  606..2225  
                  /gene="groEL"  
                  /codon\_start=1  
                  /transl\_table=1  
                  /product="groEL"  
                  /protein\_id="GI:2893007"  
                  /db\_xref="GI:2893007"  
                  /translacion="MAKEIKFSEBARAMLRGVDKILADAVKVTLPGRNVLEKKEG  
                  SPLITNDGVTIAKEILEDEPENNAGALVAEVAASKTDVAGDGTITATVLAQAMIRG  
                  LKVTAGANPMGIKIEKAIVAAVEELKASIKIOEKESIAQVAALSADEVEGOLI  
                  AEMERVGNDGVITPLEESKGFATLDVVEGMOFGRGVSPYMTIDTEKMEAVLENPYI  
                  LIDDKYSSIOELITILEVOVGSRLLIIAEDVEGALATLVVNRKRGFNNAVVA  
                  PGFDKRKAMLEDAIITGEVISEELGRELSKSTIASLGRASVVATKEHTIVEGA  
                  GDSKRIKARINQIAQLSEETSEFDRKLIQERLAKLAGVAIVKGAATETELKERL  
                  RIDBALNSTRAAVEGIVAGGTALNMVSKVAIEEGDEATGKVLRLAIEEPVQ  
                  IAOTGLGESSIIIVERLKEKPGIGFNATGEMVDMIBAGIVDPITKVRSLQNNASVA  
                  AMFLTEAVYVADPEENKGNPGMDGMM"  
BASE COUNT    771 a      430 c      553 g      555 t  
ORIGIN  
Query Match                    50.5%: Score 838.8; DB 2; Length 2309;  
Best Local Similarity 70.9%: Pred. No. 7e-149;  
Matches 1113; Conservative 0; Mismatches 457; Indels 0; Gaps 0;

QY 314 TGAAGACTAAAAATGTGACAGCAGTGTCAATCCATTTGGTATCCGTCGAGCATTTGA 373  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 905 CGAAGATTGAAAAACGTTACAGCTGGGGCTAACCATTGGGCATCCGTAAGATTGA 964  
QY 374 AACAGCAACAGCACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACTGTATCTGCNA 433  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 965 AAAACGGCTGCTGGGAGTAGAAGAAATTAACAATCTCCAAACCAATTAAGATTGA 1024  
QY 434 GGAACCTATTCGTCAGTGGCGGCTGATTCATCCAGCGCTGAAAAGTTGGAGATTAT 493  
          |||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1025 AGAATCGATTGCTCAAGTTGACAGCGATCTCGCGGCTACAGAAAGATTGGCAATTAT 1084  
QY 494 CTCGAACCTATGAGAGCCTGTGGCAGACATGATGTTGATTAACATCAAGAAATTCGAG 553  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1085 CGCAGAACCAATGCAACCCGTTGGCAGACATGCTTTATCACATTAGAAGATCGAAAG 1144  
QY 554 TATGAAAACAGAACTGGAAGTGTGAAGGATGCAATTTGACCGTGTACCTGTCTCA 613  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1145 CTTCGCAACGGAATTAGATGTTGTAAGGTATGCAATTTGACCGTGTATGATCTCC 1204  
QY 614 ATACATGCTCACAGACATGAAAAATGTTGCAACCTTGAAACCATTATCTTAT 673  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1205 ATACATGATCACAGATACAGAAAAATGGAAGCAGTGTGAAATCCATCATCTTAT 1264  
QY 674 CACGATAAAAAAGTGCACAAACATCCAGACATTTGGCAGTCTGAGTGGAGAACTCTNA 733  
          |||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1265 TACAGATAAAAAGTTTGTAGCATCCAGAAATCTGCGTATCTTAGAACAAGTAGTTCA 1324  
QY 734 AACCAACGCTCATCTACTATTAATGTCAGATGATGATGATGATGATGATGATGATGAT 793  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1325 ACAAGAGACCGGCTATTAATATTGTCAGAAAGATGTCAGAAAGCGGCTGCGCAACAT 1384  
QY 794 TGTCTGAACAAGATTCGTGTGATCTTCAATGTTGTTGCTGTCAAAAGCGCCAGATTGG 853  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1385 AGTGTCAACAACCTCTGTGTACATTCATTCATTCATTCATTCATTCATTCATTCAT 1444  
QY 854 TGATGCTGTAAAGCTATGCTTGAAGACATGCTATCTTGAAGAGTGTGATGATGATGAT 913  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1445 CGATGCTGTAAAGCGATGCTTGAAGATATGCGGATTTAACTGGCGGTGAAGTTATCTC 1504  
QY 914 AGAGATCTAGACTGTAATTAAGATGTCATCAATGACAGCCCTTGACAGGCTGTAA 973  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1505 CGAAGAACTAGAGCTGAATTAATTCGAACAATTCATTCATTCATTCATTCATTCAT 1564  
QY 974 GATTACATTTGATTAAGATGACACAGTATTTGTAAGTTTCAGGAAGTTCAAGAGCTAT 1033  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1565 AGTAGTTTAAAGAAATATACAAATGCTAGAAAGGCGCTGCGATTCATAAGCAT 1624  
QY 1034 TGCTAACGTTATGCTGATTAATGCAATTTAGAAACAACAACCTTGACTTTGACCG 1093  
          |||    |||    |||    |||    |||||    |||||    |||||    |||||    |||||  
DB 1625 TAAAGCTGCTCAACAACATCCGTCGCAATTAAGAAACATCTTGGAAATTCGACCG 1684  
QY 1094 TGAANAACCTAGAACCTTTGGCGAATTTAGCTGTGTAGCTGTATTAACAAGTAGG 1153  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1685 CGAAAAATTAACAAGAACCTTTGCAAAATTTAGCTGGGCGCTAGCGGTAAATCAAGTTGG 1744  
QY 1154 AGCTTCAACAGACAGCTTTTAAAGAAATGAACCTTGCAATGAGATGCTTAATGC 1213  
          |||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1745 TCGCGCGACAAACAGATTTGAAGAAAGCAAAATTTGCGATTGGAAGCGGCTCAATC 1804  
QY 1214 TACAGCTGACGCTGTGAAGAAAGTATGCTGTGTGTGTGAAGAGAGCACTTATAGCT 1273  
          |||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1805 TACTGCTGCGCTGTCGAGAAAGATATGATGCGCGGCTGTAGGATTAATTAAGACT 1864  
QY 1274 TATTGAAAAGTAGAGCTGTGAAGCTTGAAGGCGATGATCTAGTGAAGCTTAACATTGT 1333  
          |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
DB 1865 ATACAGCAAGTGGCTGCAATTCGAAGGCGAGGAGCAACATGCTGTGAAGAAATGCT 1924  
QY 1334 GCTTCGTGCTCTAGAAGAGCTGTATACGTCAAATTCCTTAAATGCTGGGTACGAAGCTC 1393  
          |||||    |||    |||    |||    |||    |||    |||    |||    |||    |||  
DB 1925 TCTTCGCGCAATCGAAGAGCGATGTCAAATTCGCAAACTGCTGTGTTGGAAGGCTC 1984

QY 1394 CGTAGTTATGACAGTTGAAAAACAGCCCTGCAGACAGATTTATGCTGCACAGG 1453  
 Db 1985 TATCATTTGTTGAACGCTTAAACACAGAAAACTGGGATCGGCTTCACCGGCTACTGG 2044  
 QY 1454 TGATGGGCTGATATGATTAACACAGAAATCATTTGACCTGTCAAAAGTAACAGATCAGC 1513  
 Db 2045 TGATGGGCTGATATGATTAACACAGAAATCATTTGACCTGTCAAAAGTAACACTGCTGC 2104  
 QY 1514 GCTTCAAAATGACAGCTTCTGCTAGCTACTCTTATTTTGACACAGAACGACTTCTCTAA 1573  
 Db 2105 ACTTCAAAACGACGCTTCTGCTAGCTACTCTTATTTTGACACAGAACGACTTCTCTAA 2164  
 QY 1574 TAACTGAA 1583  
 Db 2165 CAACACGAA 2174

RESULT 6  
 LOCUS BACROESL 2400 bp DNA BCT 26-APR-1993  
 DEFINITION Bacillus subtilis heat shock protein (groEL and groES) genes,  
 complete cds.  
 ACCESSION M81132  
 VERSION M81132.1 GI:143025  
 KEYWORDS groEL gene; groES gene; heat shock protein.  
 SOURCE Bacillus subtilis (Strain W168) DNA.  
 ORGANISM Bacillus subtilis  
 Bacteria; Firmicutes; Bacillus/Clostridium group;  
 Bacillus/staphylococcus group; Bacillus.  
 REFERENCE 1 (bases 1 to 2400)  
 AUTHORS Li, M. and Wong, S.-L.  
 TITLE Cloning and characterization of the groESL operon from Bacillus  
 JOURNAL Subtilis  
 MEDLINE J. Bacteriol. 174, 3981-3992 (1992)  
 FEATURES 92283753

source Location/Qualifiers  
 1..2400  
 /organism="Bacillus subtilis"  
 /strain="W168"  
 /db\_xref="taxon:1423"  
 253..289  
 /evidence=experimental  
 253..259  
 276..282  
 340..346  
 /gene="groES"  
 /note="putative"  
 340..638  
 /gene="groES"  
 354..638  
 /gene="groES"  
 /standard\_name="GroES (cpn10)"  
 /codon\_start=1  
 /transl\_table=11  
 /function="Molecular Chaperone"  
 /evidence=experimental  
 /product="heat shock protein"  
 /protein\_id="AAA22502.1"  
 /db\_xref="GI:143026"  
 /translation="MLKPLGDRVTEIYSESEKTAAGTLPDSAKEKPEQGIYVAGS  
 GRVLSGERVALVEKEDRIIRISKYAGTEVKEGTETLLRESDLAVIG"  
 669..675  
 /gene="groEL"  
 /evidence=experimental  
 669..2319  
 /gene="groEL"  
 685..2319  
 /gene="groEL"  
 /standard\_name="GroEL (cpn60)"  
 /codon\_start=1  
 /transl\_table=11  
 /function="Molecular Chaperone"  
 /product="heat shock protein"

gene  
 RBS  
 CDS

/protein\_id="AAA22503.1"  
 /db\_xref="GI:143027"  
 /translation="MAKETKFESEARRAMLRGVADALADAVKYLTPGKGRNVLEKKFG  
 SPLINDGVYTIKAELEDEAFENMGAKLAEVASKINDYAGDGTATYLAQAMIEG  
 LKRVTPAGPVRKMGROAVVAIENLEISKIPJEGKSIQNOVAISADEAVGSLI  
 AEMERVNDGVYITTEESGPTTELEVNGOPDRYASPVYVTDSDKREAVLYNDYI  
 LITDKITNIQIELPYLEVOYQGRKLLIAEDVEBALATILYKLRTPFNAYKA  
 PGGDRKRMLEDIAVLVTEGEVITTEDGLDLSQTOUGRSKVVYKENTIVEGA  
 GETDKISARVQIQAQVEETSEFDEKLOERLAKLAGVAVIKVQAATELEKRL  
 RIEDALNSTRAVVEEIVSGGTALVNVNKAAYAEABDQOTGIVLRALDEPTRO  
 IAHNAGFESVIVERLKNEDIGFMAAGVNMVMEKCTIVDPTKVTBSALQNAASVA  
 AMFLTEAVYVAAKPEENGAGAMPDMGSGMGNGMM"

BASE COUNT 757 a 475 c 584 g 584 t  
 ORIGIN

Query Match 48.6%; Score 807.4; DB 2; Length 2400;  
 Best Local Similarity 68.0%; Pred. No. 5.9e-143;  
 Matches 1126; Conservative 0; Mismatches 531; Indels 0; Gaps 0;

QY 1 GAATTCGGCTTATGCAATGCAAAAGAAATCAAAATTTTCAGCAGATGCGCTGCTCCATGG 60  
 Db 671 GAGGTCTTAAACATGCGCAAAAGAAATTAAGTTAGTGAAGAAGCTCCGCGGCAATGC 730  
 QY 61 TCGCGGAGCTGATATGTTAGCAGATACCGTCAAGTACGCTTGCTTAAGGCGCA 120  
 Db 731 TTGCGGCTGTCGATGACCTTCTGATCTCTTAAGTAACTTTAGGCCAAAGGACGCA 790  
 QY 121 ATGTGTTCTGAAAAAGCTTTGTTCTCCCTTAATTAATGACGGGCTAACAATTG 180  
 Db 791 ACGTGTTTAGAAGAAATTTGCTGTTTATCAACAAATGACGGGTACATTCG 850  
 QY 181 CTAAGAGATCGAATTAGAAGATCAATTTGAAAAACATGGGAGCAAAATTTGCTGAAG 240  
 Db 851 CTAAGAAATCGAGCTAGAGAGCGGCTTGAACAAATGGGTGCTAAGCTTGTCTGAAG 910  
 QY 241 TGGCTCTTAAACCATGATATTTGCTGCTATGGGAGCTACTGCACAGTTTGAAC 300  
 Db 911 TACCGACCAAAACAAAGACGTTGCCGCTGACCGTACAACTGCAACAGTTCTTCCGC 970  
 QY 301 AAGCCATTTTCATGAAAGGCTTAAAAATGTGACAGCGGTGTAATCCAAATGATTCG 360  
 Db 971 AAGCAATGATCGGTGAGGCTTAAACGTAACAGAGCGGTAAACCTGTAGGGGTGC 1030  
 QY 361 GTGAGGCAATGAAAGCAACAAACAGCAACAGCTTGAAGCTTGAAGCCATGCTCAAC 420  
 Db 1031 GTAAAGGATGGAACAAGCTGTAGCGGTGCGATCGAAACTTAAAGAAATTTCTAACG 1090  
 QY 421 CTGTATCTGCAAGAACTATTTGCTAGCTGCTGCGATGATATCATCGCTTGAAAG 480  
 Db 1091 CAATCGAAGGCAAGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150  
 QY 481 TTGAGAGATATATCTGAGAACTATGAGAGGCTGAGGCAACGATGCTGATTAACATG 540  
 Db 1151 TCGAAGCCTTATGCTGAGGCAATGAGAGGCTGAGGAAACGAGCGGTATACATG 1210  
 QY 541 AAGAATCTGAGGTATGAAACAGAACTGTAAGTGTGGAAGCATGCAATTTGACCGTG 600  
 Db 1211 AAGAGCTAAAGCTTCAACACAGCTGAGCTTGAAGTATGCAATTTGACCGCG 1270  
 QY 601 GTTAACTGTCTAATACATGCTGACAGACATGAAAAATGTTGCAAGCTTGAAGACC 660  
 Db 1271 GATATGCGTCTCTTACATGCTGATGCTGATGATGATGATGATGATGATGATGATG 1330  
 QY 661 CATTTATCTTAATCAGAGATAAAAGGTGCAACATGCAAGCATTTTGGCCTACTG 720  
 Db 1331 CTTACATCTTAATCAGAGATAAAAGGTGCAACATGCAAGCATTTTGGCCTACTG 1390  
 QY 721 AGGAAGTTCTTAAACCAACCGTCAATTAATCAATTTGAGATGATGATGATGATGATG 780  
 Db 1391 AGGAGGTGTTGCAAGCAAGCAACATTTGCTGATCGTGAAGATGTTGAAGGCGAAG 1450  
 QY 781 CACTTCAACCTTGTCTTGAACAGATTCGTGTAATGTTGTTGCTGTCAAG 840

D	1451	CACCTTGCACACTGTTGTGTGAACAACACTTCGGCGACATTCACAGCAGTGCTCTTAAG	1510
OY	841	CGCCAGCAATTGGTGATGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTG	900
D	1511	CTCCGTGGTTTGGGTGATGCCCGTAAAGCAAAGCTTGAAGACATGCTGTCTCTACTGGCG	1570
OY	901	GTAACGATTTNACGAGGATCTAGACACTGAAATTAAGAATGTCACMAATGACACCCTTG	960
D	1571	GAGAGTATATCACAGAAATCTTGCGCTTAGCTGAAATCTACTCAATTCCTCAATTGG	1630
OY	961	GACAGGCTGTACAGATTATACAGTTGATAAAGATAGCACAGTATTTGTTAAGGTTACAGAA	1020
D	1631	GACGGCTCTTCAAGTTGTCGTACTATAAAGAACACACACATGCTTGAAGCGCTGGCG	1690
OY	1021	GTTGAGAGCTATTGTCATACCGTATTTGCATGATTAATTCGCAATTTAGAAAACAACATT	1080
D	1691	AAACAGACAAATATTTCTCCCGTGTGACTCAATTCGCCGCTCAAGTGGAGAACCAACTT	1750
OY	1081	CTGACTTTGACCCGTGAAAAACTACAAAGACGTTTGGCAATTTAGCTGTGTGTAGCTG	1140
D	1751	CTGATTTGCAAGAGAAAAATTAACAGAGCGCTTGCTCAAACTTGCGCGCGCTACTCTG	1810
OY	1141	TTCATCAAGTATGGAGCTCCACAGAGACAGCTTTAAAAAGAAATGAACCTTGCAATTGAGG	1200
D	1811	TCTATCAAGTATGGCGCTCGACTGAACTGAGCTGAAAGACGTTAACTTGCAATCGAG	1870
OY	1201	ATGCTCTAAATGCTACACAGTGCACCCGTTGAAGAGCATGCTTGTGCTGTGGTGAACAG	1260
D	1871	ACGCGTTGAACTCAACTGCGCAGCTGTTGAAGAAAGGATGTATCCGGTGTGTACAG	1930
OY	1261	CACCTATTACGGTATATGAANAATAGCAGCTCTTGAGCTTGAAGCGGATGATGCTACTG	1320
D	1931	CGCTTTGTAAGGTATATTAACAAGTCGCTGACGTTGAAGCTGAAGGCGGATGCTTCMAACAG	1990
OY	1321	GACGTAACATGTGTGGTGTGCTGTAGAAAGCCGTGAACGTCAAATGTCTTAATGCTG	1380
D	1991	GTCATCAACATGTGCTTCTGCGCGCTTGAGAGCCAAATCCGTCNAATGTCACACACAGCTG	2050
OY	1381	GGTAGAGCGCTCCGTAGTATTTAGACAAGTTGAANAACAGCCCTGCAGCAACAGATTTA	1440
D	2051	GCCTTGAAGAGTATGCTATCGTTTAGCGCCCTCAAAAAGAAATGCGCGTAGGCTTCA	2110
OY	1441	ATGCTGCACACAGTGTAGTGGTTGATATGATTAAACAGGATCATTGACCCTGTCCAAG	1500
D	2111	ACGCTGCAACGCGGAATGGGTAAACATGCAAAAAAGATGCTGTAGCCACCAACAAAG	2170
OY	1501	TAAACAGATCAGCCCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTTTGAACAACAG	1560
D	2171	TTAACGCTGTACGCTCTTCAAAACGCTGTGCTGTAGCTGCATGTTCTTAAACAACAG	2230
OY	1561	CAGTTGTGCTAAATAAACCTGAACAGCTAGCAGCCAGCGCCAGCAATGCCAGCAGTATGG	1620
D	2231	CTGTGTCGCTGACAAAGCAGAAACAAAAGSGTGGCGGCGCAGGAATGCTGTATATGGCG	2290
OY	1621	ATCCAGAGATGATGGGTGGGATGGCGGATTAAGCGCA	1657
D	2291	GCATGGCGCGTATGGGTGAATGATTAATTAAGGGA	2327
RESULT	7		
BACHSPA			
LOCUS	BACHSPA	2525 bp	DNA
DEFINITION			RCT
ACCESSION		M4965	26-APR-1993
VERSION		M84965.1	GI:143061
KEYWORDS			Atipase; chaparonin protein; heat shock protein.
SOURCE			Bacillus subtilis (strain Mbl1) DNA.
ORGANISM			Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
REFERENCE			1 (bases 1 to 2525)
AUTHORS			Schmidt,A., Schlesswohl,M., Voelker,U., Hecker,M. and Schumann,W.

FEATURES	Source	Location/Qualifiers
JOURNAL	Cloning, sequencing, mapping and transcriptional analysis of the <i>groES</i> operon from <i>Bacillus subtilis</i>	
MEDLINE	<i>J. Bacteriol.</i> 174, 3593-3599 (1992)	
FEATURES	92283754	
gene		1..2525
CDS		/organism="Bacillus subtilis"
		/strain="MB1"
		/db_xref="taxon:1423"
		168..452
		/gene="groES"
		168..452
		/gene="groES"
		/codon_start=1
		/transl_table=1
		/product="heat shock protein"
		/protein_id="AA022530.1"
		/db_xref="GI:143062"
		/translation="MKRPIDRVIVELVESEETKASSIVPDSKKEPQSKIVAGS GRVLESGERVALEVKESDRITFFSKYAGTEVKETELILRESDLIVIG"
		499..2133
		/gene="groEL"
		499..2133
		/gene="groEL"
		/codon_start=1
		/transl_table=1
		/product="heat shock protein"
		/protein_id="AA022531.1"
		/db_xref="GI:143063"
		/translation="MAKEIKSEEARARMLRGDALADAVKYLPGKRVVLEKRG SLPTINDGVIITAKEIELEDNFENKGALIEVASKTIDVAGDGTITRVLQAQVIREG LKNVTGAPVGVKRGMEQAVVAILENLKTSKPEKESIAQVAAISADEEVSILI ALEAMERGVNDGVIITIEESKQFTTELVELEOMODRGVASYPMVTSDKMAVLDNPYI LIDYKLTINIOELIPVLEQVVOQKPLLLIAEDVEGALATLVNKLRTGFNAVAKVA PGEDRRKAMLEIDIAVLVEGVEITEDIIDLUKSQLQOLRASKVYVTKENTIVEGA GEMTKISARVTOIRAOVEERTSEPRDKLOERLAKLAGVAIVKVGAAETTELKERL RIDALNSTRAVAEEGIVSSGGTALVNVYTKVAVAEAGRAQGTINIVLALEPIRL IAHNAGESVIVERLKNKEIIGVEFNATSEVMNTEKGIIVDPKTVRSALONMAASVA AMLEITAAVVAADKREENGCGACMPDMGGMGMGM"
BASE COUNT	781 a 504 c 603 g 637 t	
ORIGIN		
Query Match	48.6%; Score 807.4; DB 2; Length 2525;	
Best Local Similarity	68.0%; Pred. No. 5.9e-143;	
Matches 1126; Conservative	0; Mismatches 531; Indels 0; Gaps 0;	
QY	1 GAATTGGGCTTCATATGGCCAAAGAAATCAATTTTCAGCGAGATGCGGTCTGCGCATGG	60
Db	485 GAGGCTTGTAAACCTGGCAAAAGAAATTAAGTTAGTGAAGAGCTGCGCGCAATGC	544
QY	61 TGGCGGAGTTCATATGTATGACAGATACCGTCAAGTAAGCTTGGTCTTAAGGGCGCA	120
Db	545 TTCCGCGGTGTGATGCACTTCTGCTGATGTGTTAAAGTACTTTAGGACCAAAAGAGCGCA	604
QY	121 ATGTGTTCTTGA AAAAGCTTTTGTTCTCCCTTAATTAATTAAGAGCGGGTAAACATTTG	180
Db	605 ACGTGTCTTGAAGAAAAAATTCGTTCTCCGTTAATCACAATAGACGGGTAAACAATCG	664
QY	181 CTTAAGAGATGGAATTAAGAAATCATTTTGA AAAACATGTGGAGCAAAATTTGGTGTGAAG	240
Db	665 CTTAAGAAATGAGCTAGAAAGACGCGTTTGA AAAACATGTGGAGCTTGTCTTCTGAG	724
QY	241 TGGCTTCTAAACCAATGATATTCCTGGTGAATGGAGACACTACTGCAACATTTTGAAC	300
Db	725 TAGCGAGCAAAAACAAGACAGCTTGCGGTGACGGTACACAACTGCAACAGTTTCTTGCGC	784
QY	301 AAGCCATTGTTCAATGAAGAGCTAA AAAATGTGACAGCAGGTGCTAATCCAAATGCTATCC	360
Db	785 AAGCAATGATCCGTAAGAGCCTTAAAAACGTAACAGCAGGCGCTAACCCCTGATAGCGTCC	844
QY	361 GTGAGGCAATGAACAGCAACAGCAGCTGTGAAGCCTTAAGAACCAATTCGCTAAC	420

Db 845 GTAAAGGATGGAACAGCTGAGCGTTGCCATGCAAAACCTTAAAGAAATTTCTTAAC 904  
 Qy 421 CTGATCTGGCAAGAGAGCTATGCTCAGGTGCTGCAGTATCATCAGCCTCGAAGAG 480  
 Db 905 CAATGGAAGCAAGAGATGCTATGCTCAGGTGCTGCAGTATCTCTGCTGATAGGAAG 964  
 Qy 481 TTGAGACAGTATCTCAGAGCTATGAGCGTGTGGCAACGATGCTGATTCATG 540  
 Db 965 TCGGAAGCCCTTATCGCTGAGAGATGAGCGGCTAGGAACGACGCGCTTATCACAAATG 1024  
 Qy 541 AAGATCTTCGAGTATGAGAACAGAACTTGAAGTGTGAGAGCATTAATTGACCGTG 600  
 Db 1025 AAGAGCTTAAGGCTTCAACAAGTGAAGTGTGAGAGCATTAATTGACCGTG 1084  
 Qy 601 GTTACCTGTCTCAATCATGCTGACAGACATGAAAAATGTTGAGACCTTGAAGACC 660  
 Db 1085 GATATGCGCTCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1144  
 Qy 661 CATTTATCTTATCAACGCTTAAAAAGTGTCAAACTTCAACATGATTTGCCACTACTTG 720  
 Db 1145 CTTACATCTTATCAACGCTTAAAAAGTGTCAAACTTCAACATGATTTGCCACTACTTG 1204  
 Qy 721 AGGAAGTCTTAAACCAACGCTTCAATCTATTTGATGATGATGATGATGATGATGATG 780  
 Db 1205 AGCAGGTTCTTACAGACAGCAACCATTCCTTCTGATCCCTGATGATGATGATGATG 1264  
 Qy 781 CACTTCCAAACCTTGTCTTGAACAAGATTCGTGATCTTCAATGATGATGATGATGATG 840  
 Db 1265 CACTTGTATCACTTGTGTCGACCAAACTTCGCGGACATTCACGAGATGATGATGATG 1324  
 Qy 841 CCGCAGAGATTTGATGATGCTGTAAGAGTATGCTTGAACATTTGCTATCTTGACAGTG 900  
 Db 1325 CTCCGCTTGTGCTGATGCTCCGCTAAAGCAATGCTTGAACATGCTGCTTACTGCTG 1384  
 Qy 901 GTACAGTGTATCAGAGATGATGAGTGAATGATGATGATGATGATGATGATGATGATG 960  
 Db 1385 GAGAGTGTATCAGAGATGATGAGTGAATGATGATGATGATGATGATGATGATGATG 1444  
 Qy 961 GACAGGCTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1445 GACGCGCTTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1504  
 Qy 1021 GTTGAAGAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 Db 1505 AAACGAGCAAAATTTCTGCGCTGATCAAAATCCGCTCAAGGAGAAACAACATTT 1564  
 Qy 1081 CTGACTTTGACCGTGAAGAACTACAAAGATTTGGGCAATTTAGCTGATGATGATG 1140  
 Db 1565 CTGATTTGACAG 1624  
 Qy 1141 TTATCAAGTATGAGAGCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
 Db 1625 TCATCAAGTATGAGAGCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1684  
 Qy 1201 ATGCTCTAAATGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
 Db 1685 ACGCTTAACTCACTGCGAGCTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1744  
 Qy 1261 CACTTATTTACGTTATTTAAAAAGTATGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 1320  
 Db 1745 CCGCTTGTAAAGTATATTAACAAAGTCCGCTGCAAGTGAAGAGAGAGAGAGAGAG 1804  
 Qy 1321 GACGTAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
 Db 1805 GTATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1864  
 Qy 1381 GGTAGAGAGCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
 Db 1865 GCGTTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1924  
 Qy 1441 ATGCTGCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
 Db 1925 ACGCTGCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1984

Qy 1501 TACACAGTACAGCGCTTCAAAATGACGCTTCTGTACTAGTCTTATTTTGACAACAGAG 1560  
 Db 1985 TTACACGCTACAGCTCTTCAAAAGCTGCGCTGTACTAGTCTTATTTTGACAACAGAG 2044  
 Qy 1561 CACTTGTCTTATTAACCTGAAACCACTACAGCGCAGCAGCATGCCAGCATGATG 1620  
 Db 2045 CTGTGTGCTGACAGCAGCAGAGAAACGCTGCGCGCAGGAAATGCTGATATGAGGCG 2104  
 Qy 1621 ATCCAGATGATGAGTGGGATGCGCGGATTAAGCCGA 1657  
 Db 2105 GCATGGCGGATGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATG 2141  
 RESULT 8  
 AB007637  
 LOCUS 17516 bp DNA BCT 13-FEB-1999  
 DEFINITION Bacillus subtilis genomic DNA containing groES to gtrR region, 48  
 ACCESSION AB007637  
 VERSION AB007637.1 GI:2521992  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bacillus subtilis (strain:Marburg 168) DNA.  
 Bacteria; Firmicutes; Bacillus/Clostridium group;  
 Bacillus/Staphylococcus group; Bacillus.  
 REFERENCE  
 AUTHORS  
 TITLE  
 Kasahara,Y., Nakai,S., Ogasawara,N., Yata,K. and Sadate,Y.  
 Sequence analysis of the groESL-cotA region of the Bacillus  
 subtilis genome, containing the restriction/modification system  
 genes  
 JOURNAL  
 DNA Res. (1997) In press  
 REFERENCE  
 AUTHORS  
 TITLE  
 Submitted (03-OCT-1997) to the DDBJ/EMBL/GenBank databases. Yoshito  
 Sadate, National Institute of Genetics, Radioisotope Center; Yata  
 1111, Mishima, Shizuoka 411, Japan (E-mail:ysadate@lab.nig.ac.jp,  
 Tel:81-0559-81-6870, Fax:81-0559-81-6870)  
 FEATURES  
 source  
 1..17516  
 /organism="Bacillus subtilis"  
 /strain="Marburg 168"  
 /db\_xref="taxon:1423"  
 /map="48 degree"  
 1..275  
 /gene="groES"  
 <1..275  
 /gene="groES"  
 /note="10 KD chaperonin (protein CPN10) (protein GroES)"  
 /codon\_start=3  
 /transl\_table=11  
 /product="chaperonin"  
 /protein\_id="BAA22746.1"  
 /db\_xref="GI:2521993"  
 /translation="VGDRVVLELVESEKTAAGVLPDASAKRQEGKIVAAAGGRVL  
 ESGERVALEVEKEDRIIFSKYAGTEVEKTEGYLLRESLIIAVIG"  
 306..1956  
 /gene="groEL"  
 306..312  
 /gene="groEL"  
 312..1956  
 /gene="groEL"  
 /note="60 KD chaperonin (protein CPN60) (GroEL protein)"  
 /codon\_start=1  
 /transl\_table=11  
 /product="chaperonin"  
 /protein\_id="BAA22747.1"  
 /db\_xref="GI:2521994"  
 /translation="MAKIKFSEERARRMLRGVDALADAVKTLGPKGRNVLEKRG  
 SLITNGVATLAKILELDAENNGAKLVAVASAKTNDVAGDTTATVLAQATIRG  
 LKNTAGANPVTGYKKGMEQAAVAIENLKEISKIRIEKESIAQVAAISAADDEGSLI  
 AEAMERVNDGVITIEESKGFTELEVVEGQFDRIYASPYMTVDSDKMEAVLDNPTI

```

LTDKKTITNIOELIPVLEQVYQGGPKLLLAEDVEGALATLVVNNKLGGTFNNAVKA
PGFDRRRKAMLEDAVLTGGEVITEDLGDLKSTQIQLQGRASKVVYTKETTTVEGA
GETDKISARVQIQAQVEETSEFDRKLEKRLAGVAVYKGAATETLELKERL
RIEDALNSTRAAVEGIVSGGTALVNVNVAVAVEAGDQGTGINIVLRALPEPIRO
IAHNAGLEGSYIVERIKNEEIGVGNATYGMVIMKEIGYIDPTKTYRSALQNNASVA
AMPLTTEAVYADKREENGCGAGMPDMGCMGCMGM"
RBS      3506..3514
/gene="ydlm"
/gene="ydlm"
3506..3900
/gene="ydlm"
3520..3900
/gene="ydlm"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22748.1"
/db_xref="GI:2521995"
/translacion="MISDNNFVSETELENNVOYLLPGCAVYIKRGYSRAHKYTTAKSEY
EKMVAAGLSGSEIALLIRKGHWIGASIPAGGIVIDIDSKQGLVGLLDAQNFCH
CIRTLMGSLFLRTNMGKRRKLNK"
RBS      5056..5061
/gene="ydlm"
5056..5282
/gene="ydlm"
5067..5282
/gene="ydlm"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22749.1"
/db_xref="GI:2521996"
/translacion="MIKFSVILGIRCSLTHITTKNTVNAKLMTYPRKQSPFFHEEK
VLXKLKFKCIGIKIMIKIRSCMGYFL"
RBS      5301..5307
/gene="ydlm"
5301..6595
/gene="ydlm"
5312..6595
/gene="ydlm"
/note="similar to modification methylase"
/codon_start=1
/transl_table=1
/protein_id="BAA22750.1"
/db_xref="GI:2521998"
/translacion="MTNFIENENKOLSLATEDENIENFYIDGTDVLRKTIIRSGSGVT
SRVPLSTODLENNKNTHELYDESWLBMKNRPNTLETSEINADIPEGCGGLSGVYE
ACRALGINRPFACDLNEALSYEKNRSPDSLNEIKHINDELGAFLIVEEORI
KDRVVKRIDFLAGPCCQGHSDLNHNRKDPNRNALMARSRVLELQPSVILVENQNG
I1HDKSGSFKEFNHNLKTOGYRDEIVLNAEKLGVQARRRYEFAKSTPVSSLNQIN
EFTYSNRPISMAISDLVENVGDIENFNTASEHSLNNRRLLEYFENNLPENNERPD
CHRLKPHSYKSYGMYWMDRPAPTITRGFGSTGGGRVHSLKRTITPHHAKRQFFP
DFPNFGDLRRROYDVIIGNAVPSKLSYLLALHQLR"
RBS      6604..7786
/gene="ydlp"
6604..6611
/gene="ydlp"
6617..7786
/gene="ydlp"
/note="probable modification methlase"
/codon_start=1
/transl_table=1
/protein_id="BAA22751.1"
/db_xref="GI:2521998"
/translacion="MKVVSLSFGSIGIGELGLHOSGHTTEIFCEYDPLAKAVLSKNFPG
VKTEDINEIRELPSCDLYAAGPCQDLSQAGKEGIDGSRSGLVKLFELIERKEHA
NRPPWILLIENVPMLRLNKGAMSYLTSYLGTYWARTVDARCEGLQRRRIYVL
LASLFDPKDVIYFSODHSEPDLDGKPSVDHNSYVYGTGELRGVGMAREAVPIKC
GSSVGIASPAVMSPYEDIVGTINIRDAERLOGEPDMNTITTEGDKIDEGARMLRY
GNADVSVKMWIGENLSOPKGSISDEGELVTKTWPSSAAGYGDKKYKYPVSKVWNT
EOTIAISEPLNHPKLPLSARALNGLFGRACNACTVNSDEFINSLERCKDQOLQV"
terminator
RBS      8255..8262
/note="stem loop"

```

```

/gene="ydlm"
8255..8460
/gene="ydlm"
8272..8460
/gene="ydlm"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22752.1"
/db_xref="GI:2521999"
/translacion="MKSEFFLIVSTYLNCKDIRLNSILVSKRVANOIALAFVPEEIM
FSCLVVAINISGNELSKR"
RBS      9697..10653
/gene="ydlr"
9697..9704
/gene="ydlr"
9712..10653
/gene="ydlr"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22753.1"
/db_xref="GI:2522000"
/translacion="MPEIKRLEDAVEYLLGNYPAGVSTSYSKYNEIRKIYSEAPLIG
ENEYVYVGTGTSRRISLNDTRFACQGNKHTMAVAFTSIPVDSDELKKEKAPLIGTE
ITSSKPCNGEERCTIPDGILFLVPGEGTVPYLAFLVPENDSELEKKEAPLIGTE
LKEPRANOSPACQENNKSAKNEDESKSYVFLDIEDGSIYFVEDKDTYIGDM
IWTASHINGSSAITRLEVELEVENLVYCKIKHKYNEBVDNSLTKFTVNIQDLISFL
DLHPNVONGSEGFVSDIVDENATSSDDLPEDENN"
RBS      10668..10674
/gene="ydlr"
join(10668..10674,10685..11716)
/gene="ydlr"
10685..11716
/gene="ydlr"
/note="weakly similar to 5-methylcytosine-specific
restriction enzyme McrB of E. Coli"
/codon_start=1
/transl_table=1
/protein_id="BAA22754.1"
/db_xref="GI:2522001"
/translacion="METSQTSDDLISLEKKKGTLPPKFSVLSIRPRRIIYGAGTCK
SNYLEREVGKIFGDNPPYFTRVFPFGYGOFGAVKPVYIKKLSGEEIESSNR
DKMENPEPMIDYOPGPGPTDVLKALKNRYTNLIIETIRNANASVPGDIPQLLD
RNKGSDEDPYTPCEDIMNTIARIGTIDEMIKLDSNFTMATNNADQVLPDITAK
RMSFEYLELEKRAVDSWKLSTRYGHNKVIMWNPFDIIRLKLKQVPEKLLGP
FFLKESELNMQNVEKRLLYLKEDVEKHNPTIDFLNASTFSELIERYDGSNDNIFFD
IDSSFEVSD"
RBS      11705..11712
/gene="ydlr"
11705..11712
/gene="ydlr"
11719..13116
/gene="ydlr"
11719..13116
/gene="ydlr"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22755.1"
/db_xref="GI:2522002"
/translacion="MDKSSKFFPEDOKYNNERYRVIGNALALKSKGILYEDSSGCL
IFNVGYISGNRVITFLPKYCNHRHDEHSKRTLEFNKLIIFKYSGLNLSRSDYFV
SELDSDVSDFMADYLLNDPSLNGYQOKRTEVEIDEGEIIIMSKTVNSITPVFSKG
VPYFYSTNENVOQDEYHLIVKIHKWLKSYFNDFGVILGFTGLEPFSKSCDGMILDY
ADFFGSYINKEIVTYVYDRDYLLKALKTAIDRENOFSKRPYLSLGTGYPHRWME
VCKTQFSEHVAWEYVKKISRPMMINPTDIEVAKKEKTELEDDIYKAEYSKNTYPLITAK
YXININPFGKLECNPGVEDTETKOLYLYLAKLEKSRGKTKRNAEPLPSSNSRTPFKVFG
SVDPDFLDIAVAVLTIVTISAQVYMLYLENTEFTSDTLKFFVSLEIKSKRKHSHVITSTL
YGNMFLTKRLSDKN"
Query Match      48.6%; Score 807.4; DB 1; Length 17516;
Best Local Similarity 68.0%; Pred. No. 4,1e-143;

```

Matches 1126; Conservative 0; Mismatches 531; Indels 0; Gaps 0;

```

QY 1 GAATGCGCTTCATATGCGAAGAAATCAATTTTCAGCAGATGCGCGCTGCCATGG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 GAGGCTTTGTAACATGGCAAAAGAAATTAATAGTTAGTAAGAAAGCTCGCGGGCATGC 367
QY 61 TGGCGGAGTGTATATGTTAGCAGATACCGTCAAGAGTGGCTTCTTAAAGGGCGCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 TTGCGCGGTGCGATGCACTGTGATGCTGTTAAAGTAACTTTAGGACCAAGAGCGCA 427
QY 121 ATGTTCTTGAAGAGCTTTGGTTCTCCCTTAATTAATTAATGAGCGGGGTACCATG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 428 ACCTGCTTAGAGAAAATTTGCGTTCTCCGTTAATCAAAATGACGGGTACAAATCG 487
QY 181 CTAAAGAGATCGAATTAAGAGATCATTTTGAACATGGAGCAAAATTTGTTCTGAAG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 488 CTAAAGAAATCGAGCTAAGAAAGACCGCTTTGAACATGGGTGCTAAGCTTGTGCTGAAG 547
QY 241 TGGCTTCTAAACCAATGATATTGCTGTGATGGAGCAGCTACTGCAACAGTTTTCACAC 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 TAGCCGAGAAACAAACGAGCTGCGCGTGACGCTACACMACCTGCACAGCTTCTGCGC 607
QY 301 AAGCCATTGCTCAAGAGCATAAATAATGTACAGAGGTGCTAATCCAAATTTGCTATCC 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 AAGCAATGATCCGTGAAGGCTTAAACGTTAAACAGAGCGGCTTAACCTGTAGCGCTGC 667
QY 361 GTGCGAGCATTTGAACAGCAACAGCAAGCTTTGAAGCCCTTGAAGCCATTGCTCAAC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 GTAAAGGATGGAACAAGCTGTAGCGGTGCGATCGAATTTAAAGAAATTTCTAAGC 727
QY 421 CTGTATCTGGCAAGAGCTATTGCTCAGGTGCGCTGCAATATCATCGCTCTGAAAAG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 CAATCGAAGCAAGAGCTATGCTCAGAGTTCTGCAATCTGCTGATGAGAGAG 787
QY 481 TTGGAGATATATCTAGAGATATGAGCGGTGGGCAACGATGAGTATTAACCATG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 TCGGAAGCCTTATCGCTGAAGCAATGAGCGCTAGGAAACGACGGCTTATACAAATCG 847
QY 541 AAGAAATCTGAGGTATGAAACAGAACTTGAATGTTGAAGGATGCAATTTGACCGTG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 AAGAGCTTAAAGGCTTCAACACTGAGCTTGAAGTGTGAAGGTATGCAATTTGACCGCG 907
QY 601 GTTACTGTCTCAATATGATGCTCAGACAGATGAAAAATGTTTGCAGACCTTGAACAC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 GATATGCGTCTCTTACATGTAAGTACTGATTAAGATGGAAGGGTCTTTCAGCAATC 967
QY 661 CATTTATCTTAATCAGAGTAAAGTGAACATCAACAGATTTTGGCCACTCTT 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 968 CTTACATCTTAATCAGACAGCAAAAAATCAACAACATTCAGAAATCTTCTCTGCTTG 1027
QY 721 AAGAACTTCTTAACCAACCGCTCATTAATGCAATGATGATGATGATGATG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1028 AACAGGTGTTGCAAGCAACCAACCTGCTTGTATGCTGATGATGATGATGATG 1087
QY 781 CACTTCCACCCCTGCTCTTGAACACAGATTCGTGACTTTCAATGTGCTGCTCAAG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1088 CACTTCTCACTCTGTGTTGAACAACTTCGCGGCACTTCAACGAGAGTGGCTTTAAAG 1147
QY 841 CCCAGAGATTTGGTATCGTCAAGATGATGATGATGATGATGATGATGATGATG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1148 CTCTCTGTTTGGTATGCTGCTTGAAGCAATGTTGAAGCAATGCTGCTTCTTACTGCGC 1207
QY 901 GTACAGTATTAACAGAGATCTAGACCTGTAATTAAGATGCTACATGACAGCCCTTG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1208 GAGAATCTATCAAGAGATCTTGGCTGACCTGAAATCTACTCAATGCTCAATTTG 1267
QY 961 GACAGCTCTCTAAGATTAAGATGTAAGATGACACAGTAAATTTTGAAGTTCAAGAA 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1268 GACGCGCTTCTAAGTTGCTGTTACTAAGAAACCAACCAATCGTTGAAGGGCTGAGG 1327
QY 1021 GTTCAAGACATTTGCTTAACCGATATGCACTGATTAATTCGCAATTAAGAAACCACTT 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1328 AAACAGCAAAATTTCTGCGCGGTGCTGACTCAAAATCCCGCTCAAGTGAAGAAACCACTT 1387

```

```

QY 1081 CTGACTTTGACCGTGAAAACTACAGAACGTTTGGCGAAATTAAGTGTGTAGCTG 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1388 CTGAATTCGACAGAGAAAAATTAACAGAGAGCTCTTGTCTAACTGCTGCGGGGTAGCTG 1447
QY 1141 TTATCAAGTAGAGCTCCACAGACAGACCTTTAAAAAGAAATGAATCTTCCATTGAGG 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1448 TCATCAAGAGCGCGCTGCGACAGTGAAGAGCCCTAAACCTTCGATGGAAG 1507
QY 1201 ATGCTCTAAATGCTACACGCGCCGCTTGAAGAGATGCTGCTGGTGGTGGAGACAG 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1508 ACCGCTTGAACCTAATCGCGGACGCTGTTGAAGAGCATGATCCGGTGTGTAGACAG 1567
QY 1261 CACTTATTCAGGTTATTAAGAAAGTACAGACCTCTTGAAGTGGAGCGATGATGCTAC 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1568 CGCTTGAACGATATTAACAAAGCTGCTGACATGTAAGTGAAGGAGATGCTCAAAACAG 1627
QY 1321 GAGCTAACATTTGCTTCTGCTGTAGAAAGCTGCTAGCTCAATTTGCTTTAAATGCTG 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1628 GTATCAACATCGTGTGCGCGCTTGAAGAGCCAAATCCCTCAAAATCGCACACAGCTG 1687
QY 1381 GGTACGAGGCTCCGTAGTATGACAAAGTTGAACAAACAGCCCTGCAGAGAACGATTTA 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1688 GCCTTGAAGATCTGTATGCTGTAGAGCGCTCAAAACAGAAATCGCGTGTAGCTTCA 1747
QY 1441 ATGCTCAACAGCTGAGTGGGTGATATGATTAACAGAGAAATCAATGCTGTCAAG 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1748 ACCGCTCAACTGCGAATGGGTAAACATGATGCAAAAGATATCGTTGACCAACAAAG 1807
QY 1501 TAACAGATCAGCGCTTCAAAATGCAAGCTTCTAGCTACTTATTTTGAACAGAGAG 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1808 TTACAGCTCAGCTTCAAAACGCTGCGCTGTACTGCAAGTCTTAAACAGAGAG 1867
QY 1561 CAGTTGTTGCTAATTAACCTGACAGCTACGCCAGCGCAGCATGCCAGCAGATGAG 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1868 CTGTTGCTGTCAGAACCCGAAAGAAACGGTGGCGGCGAGCAATGCTGATATGAGCG 1927
QY 1621 ATTCAGAAATGATGGTGGATGGGCGGATTAACCGCA 1657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1928 GCATGGCGGTATGGGTGAATGATGATTAAGAGGA 1964

```

RESULT 9

BSUB0004 213190 bp DNA BCT 26-NOV-1997

LOCUS Bacillus subtilis complete genome (section 4 of 21): from 600701 to 813890.

DEFINITION 299107.1 AL009126

ACCESSION 299107.1 GI:2632866

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Bacillus subtilis.  
Bacteria: Firmicutes; Bacillus/Clostridium group;  
Bacillus/Staphylococcus group; Bacillus.  
1 (bases 1 to 213190)

Kunst, F., Ogasawara, N., Moser, I., Albertini, A.M., Alloni, G.,  
Azevedo, V., Berleto, M.G., Bessières, P., Bolotin, A., Borcher, S.,  
Boriss, R., Boursier, L., Brans, A., Bratu, M., Brigneau, S.C.,  
Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V.,  
Carter, N.M., Choi, S.K., Codani, J.J., Comerion, I.F., Cummings, N.J.,  
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,  
Emmerson, P.T., Enlian, K.D., Errington, J., Fabret, C., Ferrari, E.,  
Fougeret, D., Fritz, C., Fujita, M., Fujita, Y., Funa, S., Gallizzi, A.,  
Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J.,  
Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Haeck, J., Harwood, C.R.,  
Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,  
Itaya, M., Jones, L., Joris, B., Katamata, D., Kasahara, Y.,  
Klaert, Blanchard, M., Klein, C., Kobayashi, Y., Koester, P.,  
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,  
Larindols, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A.,  
Masuda, S., Maue, C., Medigue, C., Medina, N., Mellado, R.P.,  
Mizuno, M., Mostl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,  
Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,

Portetelle, D., Portwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Seror, P., Shin, B.S., Soldo, B., Sorokin, A., Taccioni, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Varti, A., Wambolt, R., Wedler, E., Wedler, H., Weltenegger, T., Winters, P., Wipal, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zmistein, E., Yoshikawa, H. and Danchin, A.

The complete genome sequence of the gram-positive bacterium *Bacillus subtilis* Nature 390 (6657), 249-256 (1997)

98044033

2 (bases 1 to 213190)

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur, Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

## FEATURES

## source

1. 213190

/organism="Bacillus subtilis"

/strain="168"

/db\_xref="taxon:1423"

33. 740

/gene="ydfs"

33. 740

/gene="ydfs"

/function="unknown"

/note="similar to hypothetical proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12373.1"

/db\_xref="GI:2632867"

/db\_xref="SPTREMBL:P96697"

/translation="MTELEVITRTVASFGLLIAERILKQTSQMTIPFIAITILG ATRAGLATNTSTIKPHMAISFSLFVLTPLISPLSIKNRKLRFAGDPVLIQNKI LESNMKMRVLTLDYLNQOLKEKIEFNIIEVLFILITNGQLVLRKPOFRRVTKODIM IAVNOCORPLIELIMDGEIENNLIKONRILFESWLBELRKDITKRETVAVILGNCD IYVDYKDHISVPMKE"

complement(755. 1186)

/gene="ydit"

/complement(755. 1186)

/function="unknown"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12374.1"

/db\_xref="GI:2632868"

/db\_xref="SPTREMBL:P96698"

/translation="MDEKIRKWLITNEYKQSDFTNVLTKAPHPFDESEASTFY DFYODEYVNIIVEMPGYEEELITLSTKLTOLLIGITLTPVPALMEVLYRENYEGBI ERIQPEAEETHTLOIOLNGLHLSYRQVTVAFNKG"

complement(1199. 1441)

/gene="ydgA"

/complement(1199. 1441)

/gene="ydgA"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12375.1"

/db\_xref="GI:2632869"

/db\_xref="SPTREMBL:P96699"

/translation="MPYOINIANIKINGTVONGNIDYGPVHNSHTANSKYFGANFSL GDSPTSSLTNTGNDISVSDQDQIGNPAISNQI"

complement(1455. 1727)

/gene="ydgB"

## CDS

complement(1455. 1727)

/gene="ydgB"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12376.1"

/db\_xref="GI:2632870"

/db\_xref="SPTREMBL:P96700"

/translation="MPSYINLYKRNISISGNSITIGFAAVNSPTNNQKSGTSS FQDTPRESYEMENFLNDPVNDQTSIGNSDTSVNNAPPIAPPILD"

2026. 2613

/gene="ydgC"

2026. 2613

/gene="ydgC"

/function="unknown"

/note="similar to hypothetical proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12377.1"

/db\_xref="GI:2632871"

/db\_xref="SPTREMBL:P96701"

/translation="MKSQVESRRRLKAAANFSGRFDADAVSELYKAGCTQSPF YLYFQSKFATFAELITDFHSVRKRLTESLLELNGLTNEDYSKRVLLAVETVQFLQED KDLTKIGFLNPEAKOMKDLAVLKENLEAEQRLGFHSELDMETVAECLVGMTEHL TESFLITGIDKDPASLAEVVNLITYGMLPGNDVR"

join(2610. 2954,3074. 3091)

/gene="ydgD"

2610. 2954

/gene="ydgD"

/function="unknown"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12378.1"

/db\_xref="GI:2632872"

/db\_xref="SPTREMBL:P96702"

/translation="MISIMKVSIAVFMLAGTITKRVFPQVEHNRHYQPLMFLVY TCSILAGALAMTAGIWNRYAIGAGVLFVLMAGALHAFMRARQSVIWAIGAMICL IVSIIIMSGSYT"

3074. 3091

/gene="ydgD"

3117. 3590

/gene="ydgE"

3117. 3590

/gene="ydgE"

/function="unknown"

/note="similar to hypothetical proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12379.1"

/db\_xref="GI:2632873"

/db\_xref="SPTREMBL:P96703"

/translation="MAFTLEDMTDEEERAGNSYONFAKONITSGTWEKAEKSE OAYENMIPNGRDSNNHYFMNITNDQGRKMGMILVADPLHPQKEAFIYSGLYEARG KGIAQLALQTLIDRRARELGAERLALHFAHNETAVVLYQMGVAMNIRRKQLC"

complement(3705. 3722)

/gene="expZ"

complement(3705. 3722)

/gene="expZ"

/complement(3705. 3722)

/gene="expZ"

/complement(3750. 5393)

/gene="expZ"

/complement(3750. 5393)

/gene="expZ"

/codon\_start=1

/transl\_table=11

/product="ATP-binding transport protein"

/protein\_id="CAB12380.1"

/db\_xref="GI:2632874"

/db\_xref="SWISS-PROT:P39115"

/translation="MKEIVLTNVSYEVKQDQTVFKHVNASVOOGDIIGITCKGACKS TIIHLIHNDLAPAGOGILRKDITLALVEQTAAYSPADQPAKRIILEKHVLYRPH QLSGGERLKRLAKGLSEADLLILDEPTNHLDEKSLIQQLKHTNGVILVSDR QYFDEAKTKIWSLEDQTLIEFGNYSGYMKERKRLTQDREYKQKVERTEAADR GLASWSERKHAQSTKKEGFEKVRVAKRTDAQIKSKOKILEKLEKAKAEPTPEYT VRSIDTHTKGRFLLEVQNVTKAFGRRLTFKNNANFTIQHGEKVALIIGPGSKTTL

## terminator

complement(1455. 1727)

/gene="ydgB"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12374.1"

/db\_xref="GI:2632868"

/db\_xref="SPTREMBL:P96698"

/translation="MDEKIRKWLITNEYKQSDFTNVLTKAPHPFDESEASTFY DFYODEYVNIIVEMPGYEEELITLSTKLTOLLIGITLTPVPALMEVLYRENYEGBI ERIQPEAEETHTLOIOLNGLHLSYRQVTVAFNKG"

complement(1199. 1441)

/gene="ydgA"

/complement(1199. 1441)

/gene="ydgA"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12375.1"

/db\_xref="GI:2632869"

/db\_xref="SPTREMBL:P96699"

/translation="MPYOINIANIKINGTVONGNIDYGPVHNSHTANSKYFGANFSL GDSPTSSLTNTGNDISVSDQDQIGNPAISNQI"

complement(1455. 1727)

/gene="ydgB"

## gene

complement(1455. 1727)

/gene="ydgB"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12375.1"

/db\_xref="GI:2632869"

/db\_xref="SPTREMBL:P96699"

/translation="MPYOINIANIKINGTVONGNIDYGPVHNSHTANSKYFGANFSL GDSPTSSLTNTGNDISVSDQDQIGNPAISNQI"

complement(1455. 1727)

/gene="ydgB"

## CDS

complement(1455. 1727)

/gene="ydgB"

/function="unknown"

/note="similar to hypothetical proteins from B. subtilis"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB12375.1"

/db\_xref="GI:2632869"

/db\_xref="SPTREMBL:P96699"

/translation="MPYOINIANIKINGTVONGNIDYGPVHNSHTANSKYFGANFSL GDSPTSSLTNTGNDISVSDQDQIGNPAISNQI"

complement(1455. 1727)

/gene="ydgB"

```

NITLQETAEKSWVSPSANIGYLTQEVFDLPLEQTPBELFENEFKANGHVNLMRH
LQFTLAOWTEPIKHMSEGRVXIKLMAVILKEKDVILIDEPNHLDPREDEETLS
QVSGTLAVSHDRYLEKTNKRLVSNNGIKOLNDVDYSENERREILKLETEROE
VIGKLSFMTPNKGYKELDOANELTKRIKEIDHODKD"
      complement(5428..5450)
      /gene="ydgf"
      /complement(5428..7089)
      /gene="ydgf"
      /complement(5678..5703)
      /gene="ydgf"
      /complement(5713..7089)
      /gene="ydgf"
      /function="unknown"
      /note="similar to amino acid ABC transporter (permease)"
      /codon_start=1
      /transl_table=11
      /protein_id="CAB12381.1"
      /db_xref="GI:2632875"
      /db_xref="SWISS-PROT:P96704"
      /translation="MTDWTNDINQDTLQRLKRNHQLIAGIGTGLFGSGKS
      /HFAGPSILFAVMITGILICFLIMRSLGELLNLYHSFVDPVDYLGDMAFITGM
      YWFCISIMADLTAVGLTYOYMLPGVPMVGLLAILILIMNLATYKLFGELEFM
      ALIKVIALILVIGLVMIFKGFSTSGVSSFTNLMSHGLEPNMGHGLISFOYVF
      APVGEIVGTAGTEMPEKVPKAIINNPVPLVLFYIGALLVISTIPWDLINSES
      PFVOVFVAVGIVGAASIIINFVLVISAASONSVAVSTSHMVISLAKDHNPBSMAKLT
Query Match      48.6%; Score 807.4; DB 2: Length 213190;
Best Local Similarity 68.0%; Pred. No. 2,6e-143;
Matches 1126; Conservative 0; Mismatches 531; Indels 0; Gaps 0;
      gene
      terminator
      CDS
QY      1 GAATTCGGCTTCAATATGCAAAAGAAATTTTCACGATGCGCGCTGCCATG 60
DB      49236 GAGGCTCTTAAACATGCAAGAAAGAAATTAAGTTAGTAGAAGAGCTCCCGGCAATGC 49295
QY      61 TCGCGGAGTGTATATGTATGATACGATACCGTCAAGTACGCTTGTCTTAAAGGCGCA 120
DB      49296 TTCGGCGTGCATGACACTTGTGATGCTGTTAAAGTAACTTTAGACCAAAAGAGCGCA 49355
QY      121 ATGTTGTTTGAAGAAAGCTTTTGGTTCCTTAATTAATGAAGGGGTAACCATG 180
DB      49356 ACGTGGTCTTACAGAAAATAATTCGGTTCCTGCTTAACAATAAGCGGTGTAACATG 49415
QY      181 CTAAAGAGATGCAATTAAGAGATCATTTGAAACATGGGAGCAAAATGCTCTGAAG 240
DB      49416 CTAAAGAAATCGAGCTAGAGAGCGCTTGAAGAACATGGGTGCTTAAGCTTGTGCTGAAG 49475
QY      241 TGGCTTCTTAAACCAATGATATTTGCTGTATGGAGCACTACTGCAACAGTTTGTACAC 300
DB      49476 TAGCCAGCAAAACAAACGAGCTTGGCGGTGAGCGTACACAACTGCAACAGTTTGTGCGC 49535
QY      301 AAGCATTTGTTATGAGAGCTAATAAATGTACACAGCGTGTCTAATTCGAATTTGATCC 360
DB      49536 AAGCATGATTCGTCGTAAGCCCTTAAAGAGCAAGCGGCTTAACCTGTAGGCGTGC 49595
QY      361 GTGAGGAGCTTGAAGACAGCAACGAGCTGTGAACCTTGAAGCCATTCGCTCAAC 420
DB      49596 GTAAAGGAGTGAACAGCTGTAGCGGTGCGATCGAATAAAGAAATTTCTTAAC 49655
QY      421 CTGTATCTGCAAGAGAGCTATTTGCTAGGTGCTGACATATCAACGCTTGAAGAAAG 480
DB      49656 CAATGGAAGCGCAAGAGCTATTCGCTCAAGTTCGTCGATCTCTGCTGTATGAGGAAG 49715
QY      481 TTGGAAGATATATCTAGAAGCTATGAGAGCTGTGGGCAACGANTGGTGTATTCATCAG 540
DB      49716 TCGGAAGCTTATCGCTGTAAGCAATGAGCGGTAGGAACGACGGCTTATCAACATG 49775
QY      541 AAGATCTTCGAGCTATGGAACAGACTTGAAGTGTGTAAGGCATGCATTTGACCGTG 600
DB      49776 AAGAGCTTAAAGGCTTCAACAGCTGAGCTTGAAGTGTGTAAGGCATGCATTTGACCGTG 49835
QY      601 GTTACCTGTCTCAATATAGTGCACAGCAATGAAGAAATGTTGACAGCTTGAAGAAC 660
DB      49836 GATATGCGTCTCTTACATGTAAGTACTGATTAAGATGGAACGCTTCTTGAACAATC 49895

```

```

QY      661 CATTATCTTAATACGAGATAAAAAGTGTCAACATCCAGACATTTTCCACTACTTG 720
DB      49896 CTACATCTTAATACAGAGCAAAAAAATATCAAAACATTCAGAAATCTCTGCTGTTG 49955
QY      721 AGCAAGTCTTAAACCAACCGCTTACTCATTTATTCAGATGATGATGATGGAAG 780
DB      49956 AGCAGCTTGTTCAGCAAGGCAACCATTCCTTGATCGCTGAGATGTTGAAGGCGAAG 50015
QY      781 CACTTCCAAACCTTGTCTGTGAACAAGATTCGTGTACTTTCATTTGATGCTGTCAAG 840
DB      50016 CACTTGTACACTTGTGTGTGAACAACCTTCCGCAATTCACGACGCTGTGAAG 50075
QY      841 CCGCAGATTTGGTGTATGTCGTGAAGCTATTCGTTGAAGCATTTGCTATCTGACAGTG 900
DB      50076 CTCCTGTTCCGCTGTATTCGCGCTTAAGCAATGCTTGAAGCATTCGCTGTCTACTGCGC 50135
QY      901 GTACAGTGTATACAGAGATCTAGACCTTGAATTTAAAGATGCTACAAATGACAGCCCTTG 960
DB      50136 GAGAACGTCATCACAGAAAGATCTTGGCTTGACCTGGAATCTCAAAATTCGAATTTGG 50195
QY      961 GACAGCTGTATGATATACAGTGTGAAGATATGACAGATATTTGTTGAAGTTCAAGAA 1020
DB      50196 GACGCGCTTAAAGTGTGTCTTCTTAAGAAAAACAAACAATCGTTGAAGGCGCTGCGC 50255
QY      1021 GTTCAGAGCTATTCCTTAACCGGATTCACATGATTAATCGCAATTAGAAACAACACTT 1080
DB      50256 AAMCACAACAATAATTTGCGCGGTGTACATTCGCGCTCAAGTGAAGAAACAACACTT 50315
QY      1081 CTGACTTTGACGCTGAAGAAATCTACAAAGACTTTGGCGAAATTAAGTGTGTGATGCTG 1140
DB      50316 CTGAATTCGACAGAGAAATAATTAAGAGAGCTTCTTAACCTTCTGCGCGGTGATGCTG 50375
QY      1141 TTATCAAGTAGAGAGCTCAACAGAGACAGCTTTAAAGAAATGAACCTTCGATGAAG 1200
DB      50376 TCATCAAGTTCGCGCTGCGACTGAAACTGAGCTGAAGAGCGCTAAACTTCGATGAAG 50435
QY      1201 ATGCTCTAAATGCTACAGCTGACCGCTTGAAGAGCTGTAGCTGATGTTGATGGAACAG 1260
DB      50436 ACGCTTGAAGCTCAACCTCGCGAGCTGTTGAAGAAAGCATCGATCCGGTGGTGAACAG 50495
QY      1261 CACTTTTACGCTTATTTGAAGAAAGTACAGCTTTCAGCTTGAAGGCGATGATGCTACTG 1320
DB      50496 CGCTTGTAAAGCTATTAACAAAGTGTGCACTTGAACCTGAAGGCGATGCTCAAAACAG 50555
QY      1321 GAGTAAACATTTGCTGCTGCTCTAGAAAGCGCTGTAGCTCAAAATTCGTTAAATGCTG 1380
DB      50556 GATACAAACATTCGCTGCTGCTGCGCGCTTGAAGAGCCATTCGTCAAATCGCACACAGCTG 50615
QY      1381 GGTACGAGGCTCCGTAATTATTTAGCAAGTTGAAAAACAGCCCTGACGAGACAGATTTTA 1440
DB      50616 GCCTTGAAGATCTGTCATCGTTGAGCGCTCAAAAACAGAAATCGCGCGTAGAGTTCA 50675
QY      1441 ATGCTCCAGCAGGTGAGTGGGTGATATGATTAATAACAGCAATCATTCACCTGTCAAG 1500
DB      50676 ACGTTCGCAACTGCGCAATGTGAATGATGCAAAAAGATATCGTTGACCCCAACAAAAG 50735
QY      1501 TAAACAGATCAGCGCTTCAAAATGACGCTTGTAGCTAGTGTCTTATTTGCAACAAGAG 1560
DB      50736 TTACACGCTCAGCTCTTCAAAACGCTGCTGTGATGCTCAAAATGTTCTTAACAACAGAG 50795
QY      1561 CAGTTTGTCTTAATAAACCCTGAACAGCTACCCAGCGCCAGCAATGCCAGCAGGATATG 1620
DB      50796 CTGTTGTGCTACAAAGCCAGAAAGAAACGTTGGGCGCCAGAGATGCTGTATTTGGCG 50855
QY      1621 ATCCAGGAATGATGGGTGAGTGGCGGATTAAGCGCA 1657
DB      50856 GCATGGCGGTATGGGTGAATGATTAATAAGGA 50892

```

RESULT 10  
BACROEL 2190 bp DNA BCT 03-FEB-1999





Db	1515	CTGAATTCGACAGAGAAAAATTACAGAGGCGTCTTGCTAACTGCTGGCGGCGTACG	1574
Qy	1141	TTATCAAAAGTAGAGGCTCCACACAGACAGCTTTAAAGAAATGAAACTTCGATTTGAGG	1200
Db	1575	TCATCAAAAGTGGGGGCTGCACACTGAAACCTAGCTGGAAGAGCGTAAACTTCGATCGAAG	1634
Qy	1201	ATGCTCTAAATGCTACACAGTGCAGCCGTTGANAAGATTCGTTGCTGGTGGACAG	1260
Db	1635	ACGGCTTGAACTCAACCTCGGGGACGCTTGAAACAAAGCATCGTATCCGGTGGTGTACAG	1694
Qy	1261	CACATTATACCGTAAATTGAAAAAAGTAGACCTCTAGCTTAGAGGCGCATGATGCTACG	1320
Db	1695	CGCTTTGTAACGTTATATACAAAGATCGCTGCACTTAACCTGAGGCGATGCTCAAAAG	1754
Qy	1321	GACGTAACATTTGCTCTCGTGCTTAGAAGAGCGCTTACGTCAAAATTGCTTTAAATGCTG	1380
Db	1755	GTAATCAACATGCTGCTTCGGGGCGCTTGAAGAGCCATCGCTCAAAATCCGACACACGCTG	1814
Qy	1381	GGTACGAGGCTCCGTAGTATTGACAAAGTTGAAAAACAGCCCTGCAGGACAGGATTTA	1440
Db	1815	GCCTTGAAGCATCTGTCATGTTGAGCGGCTCAAAAACGAAAGAAATCGGCGTAGCTTCA	1874
Qy	1441	ATGTCGCAACAGGTGAGTGGGCTGATATGATTAAACAGAGATCAATTAGCCTGTGCAAG	1500
Db	1875	ACGGTCAACAGCGAGATGGGTAAACATATGTCAAAAGAAAGTATCGTTGACCCACAAAG	1934
Qy	1501	TAAACAGCATGAGCCCTCAAAATAGCAGCTTCCTAGCTAGTCTTTATTTTGAACAACAGAG	1560
Db	1935	TTACACGCTCAGCTCTTCAAAAGCGTGCCTGTAGCTCAATGTTCTTTAACACAGAG	1994
Qy	1561	CAGTTGTTGCTAATAAATCAGAACAGCTAACGCAAGCGCCAGCAATGCGCAGGTATAGG	1620
Db	1995	CTGTTTGCCTGACACACGAGAAGAAAGGTGGCGGCGCAGAGATGCGTGATATGGCG	2054
Qy	1621	ATCCAGCAATGATGGGTGGGATGGCGGATTAAGCCAAATT	1660
Db	2055	GCATGGGGGATGAGGTGAGTATGATTAATAAGATTAAT	2094

D	1084	GGCCTCGCCTTAAAGTTGTTGTGAACAAAAGACAACACAGATCGTTGACGAAACGGAG	1143
Oy	1021	GTTTAGAAGCATATTGCCTAACCGATTTGGCACGTATTATAATGCATTTAGAACAACAATT	1080
D	1144	ATGCAAGCTGAATATCTCTAACCCGTTTAAACCAAATTCAAACCGCAAAATTCGAAGAACACTT	1203
Oy	1081	CTGACTTTGACCCGTAAGAAACCTCAAGAAACGTTTGCGSAAATTTACGTGGTGGTGAAGCTG	1140
D	1204	CTGATTTTCGATTAAAGAAAACTTCAAGAGCGCCTACTTAACCTAGCTGGCGGTGGTGGCTG	1263
Oy	1141	TTCATCAAGTAGAGAGCTCCACACAGACAGCTTTAAAGAAATGAAACTTCCTATTGGAG	1200
D	1264	TGCTAAAGATTGGTGGCGGCTACTGAAACCTGAATGAAGAGGTTAACTTCCATTCGAAG	1323
Oy	1201	ATGCTCTAAATGCTACAGCTGCACGCCGTTGAGAGAGATGCTTGGCTGGTGGTGAACAG	1260
D	1324	ACGCCCTTAACTCAACACAGCTGACGCCGTTGGAAGAAAGAAATTTGTCGTGGTGGTACAG	1383
Oy	1261	CACATTATTACGGTTATTGAAAAAGTAGACAGCTTTAGCTTAGAGGCCATGATGCTACTG	1320
D	1384	CCCCGTTTAAAGCTCCCTGAGCGTGGTCCGACAGCATTTAAATTAAGGTGACGAAGCAACAG	1443
Oy	1321	GACCTAACATTTGTCCTCGCTGCTGTAGAAGAGCCTGTAGCTAAATTTGCTTAAATVGGTGG	1380
D	1444	GTTGTTAAACATCTCTTTCTTGTCGCTCGAAGAGCCAGCTTGTCTAAATCTCTGAAACGCTG	1503
Oy	1381	GGTAGCAAGAGCTCGTAGTTATTGACAAAGTTGAAAAACACCCCTCGAGNACAGATTTTA	1440
D	1504	GACTTTGAAGAGCTCATCATCTGTTGACGCCCTGAAAGAGTGA AAAAGTCGSCATCGGCTTTTA	1563
Oy	1441	ATGCTGCAACAGGTGAGTGGGTGATATGATTTAAACAGAGAAATCATTCAGCCTGTCAAAG	1500
D	1564	ACGAGCTACTGGGGAATATGTTGAATATG-TTGAAACAGAAATCTTTGCCACCAAGAAAG	1622
Oy	1501	TAAACAGCATAGGCCCTCAAAATGACAGCTTTCTGTACTAGTCTTATTTTGACAACAGAG	1560
D	1623	TAAACAGCTCGGCCCTTCAAAATGACAAGCTCTGTATCGCTATGTTCTTCAACACAGAG	1682
Oy	1561	CAGTTTGTCTAATTAACCTGTAACACG	1587
D	1683	CTGTTGTTGCGACCTTCCAGAAAGAAG	1709
RESULT 12			
AP001508			
LOCUS	AB001508	296950 bp	DNA
DEFINITION	Bacillus halodurans genomic DNA, section 2/14.	BC1	03-AUG-2000
ACCESSION	AP001508	BA000004	
VERSION	AP001508.1	GI:10172890	
KEYWORDS	.		
SOURCE	Bacillus halodurans DNA.		
ORGANISM	Bacillus halodurans		
	Bacteria; Firmicutes; Bacillus/Clostridium group;		
	Bacillus/Staphylococcus group; Bacillus.		
REFERENCE	1 (sites)		
AUTHORS	Takami,H., Nakasone,K., Hirama,C., Takaki,Y., Masui,N., Fuji,F., Nakamura,Y. and Inoue,A.		
TITLE	An improved physical and genetic map of the genome of alkaliphilic		
JOURNAL	Bacillus sp. C-125		
MEDLINE	Extremophiles 3 (1), 21-28 (1999)		
REFERENCE	99184645		
AUTHORS	2 (sites)		
	Takami,H., Nakasone,K., Ogasawara,N., Hirama,C., Nakamura,Y.,		
TITLE	Mesui,N., Fuji,F., Takaki,Y., Inoue,A. and Horikoshi,K.		
JOURNAL	Sequencing of three lambda clones from the genome of alkaliphilic		
MEDLINE	Bacillus sp. strain C-125		
REFERENCE	Extremophiles 3 (1), 29-34 (1999)		
AUTHORS	99184646		
	3 (sites)		
	Takami,H., Takaki,Y., Nakasone,K., Hirama,C., Inoue,A. and		
	Horikoshi,K.		
TITLE	Sequence analysis of a 32-kb region including the major ribosomal		

JOURNAL	protein gene clusters from alkaliphilic <i>Bacillus</i> sp. strain C-125
MEDLINE	
REFERENCE	Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)
AUTHORS	
TITLE	4 (sites)
JOURNAL	Takami,H., Masui,N., Nakasone,K. and Horikoshi,K.
MEDLINE	
REFERENCE	Replication origin region of the chromosome of alkaliphilic
AUTHORS	<i>Bacillus halodurans</i> C-125
TITLE	Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)
JOURNAL	99356711
MEDLINE	5 (sites)
REFERENCE	Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G.,
AUTHORS	Sasaki,R., Hirama,C., Fuji,F. and Masui,N.
TITLE	Genetic analysis of the chromosome of alkaliphilic <i>Bacillus</i>
JOURNAL	<i>halodurans</i> C-125
MEDLINE	Extremophiles 3 (3), 227-233 (1999)
REFERENCE	99411980
AUTHORS	6 (sites)
TITLE	Takami,H.
JOURNAL	Genome analysis of facultatively alkaliphilic <i>Bacillus halodurans</i>
MEDLINE	C-125
REFERENCE	(In) Extremophiles in deep-sea environments (Ed.):
AUTHORS	. HORIKOSHI, K. TSUJII;
TITLE	: 249-284; Springer-Verlag (1999)
JOURNAL	7 (sites)
MEDLINE	Takami,H. and Horikoshi,K.
REFERENCE	Reidentification of facultatively alkaliphilic <i>Bacillus</i> sp. C-125
AUTHORS	to <i>Bacillus halodurans</i>
TITLE	Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
JOURNAL	8 (sites)
MEDLINE	Takami,H. and Horikoshi,K.
REFERENCE	Analysis of the genome of an alkaliphilic <i>Bacillus</i> strain from an
AUTHORS	Industrial point of view
TITLE	Extremophiles 4 (2), 99-108 (2000)
JOURNAL	20263314
MEDLINE	9 (bases 1 to 296950)
REFERENCE	Takami,H., Nakasone,K. and Takaki,Y.
AUTHORS	Direct Submission
TITLE	Submitted (22-Mar-2000) to the DDBJ/EMBL/Genbank databases. Hideto
JOURNAL	Takami, Japan Marine Science and Technology Center, Deep-sea
MEDLINE	Microorganisms Research Group: 2-15 Natsushima, Yokosuka, Kanagawa
REFERENCE	237-0061. Japan (E-mail:takahide@jamstec.go.jp
AUTHORS	URI:http://www.jamstec.go.jp/jamstec/e/bio/deepstar/research.html,
TITLE	Tel:81-468-67-3895, Fax:81-468-66-6364)
JOURNAL	Location/Qualifiers
MEDLINE	1..296950
REFERENCE	/organism=" <i>Bacillus halodurans</i> "
AUTHORS	/db_xref="taxon:86665"
TITLE	/note="alkaliphile"
JOURNAL	complement(427..615)
MEDLINE	/gene="BH0278"
REFERENCE	complement(427..615)
AUTHORS	/gene="BH0278"
TITLE	/note="unknown"
JOURNAL	/codon_start=1
MEDLINE	/transl_table=1
REFERENCE	/protein_id="BAB03997.1"
AUTHORS	/db_xref="GI:10172891"
TITLE	/translation="MRKMGILTVASIFVLACNADENETESVDONAEVEVTEDELE
JOURNAL	QDOEEFDVYEETDEEE"
MEDLINE	798..1163
REFERENCE	/gene="BH0279"
AUTHORS	798..1163
TITLE	/gene="BH0279"
JOURNAL	/note="unknown"
MEDLINE	/codon_start=1
REFERENCE	/transl_table=1
AUTHORS	/protein_id="BAB03998.1"
TITLE	/db_xref="GI:10172892"
JOURNAL	/translation="MKKIFPSVAATFLAASQAALAAPEPVSSEOGVSNAADVTPPE
MEDLINE	TVAAOSTVMISRGITVTPTDTRKTAYSNYIHRDSPHNSAALGNDRAVSPLGEGATT
REFERENCE	YASIRGNKNIPATYRFTY"
AUTHORS	1334..3475
TITLE	/gene="BH0280"

gene 1334. 3475  
/gene="BH0280"  
/note="unknown"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB03999.1"  
/db\_xref="GI:10172893"  
/translation="MHWSAAAYADFAVITPRIDLLGPDIVLSYLEEATAHQYNILR  
DVQVLGEGKQITHTTFFALLTEKSPLEAIPRHGSRLTSEDTRGDILISRRKKEDEV  
BEIKG1Q1LSQTHTRTPRPLHPQPNVSPYATQVEQAGTVTEQESATLADLFENY  
LBDYVEQLPLADFAVOTEPESHVSTFEDPLPRKTLAMTILGSGTGMXYLIMRKNK  
TVAIYSMHSEFSFKTMRRLIAGCGVAAVYLLPLLPAFLPMQORRYFLPAAVYTS  
WPLLILASIFLFVYSSRSSLSIESIKNNYTKFLFLYSQAKYVALVSYVFLQTA  
DKULTMEERNMTDMAFGKDKGYEPHISYIDDEENMQESGLMOTDISYESLYPAL  
NNAAGLYIDAVLYEEEMLMATVEYSQEHRSIRVNPNYLQAPYVDENNRKIALOEER  
TVNRVLVEYQKDKQEDILAEFKEDYAAWAMPEQVQESYSHLOEQFDTILMASQEE  
YVSFNLYPEQNGNLIDPPILOVLENNSPYMERPFPGSVGNPLPKVILDMWAST  
VSHYILPLRDLQDMDNAHLYTVNQAKDKDSEIQRALTDITTFVLTATLAFMIYQ  
VSHLEPQAKKFKFLRLRFGHSLFRATRYLFWTLATWVILGIALYRHSQTYLAIY  
VVLFLTEVYVSVLMRLQKNKYSVLKGE"

CDS 3479. 4120  
/gene="BH0281"  
3479. 4120  
/gene="BH0281"  
/codon\_start=1  
/transl\_table=11  
/product="ABC transporter (ATP-binding protein)"  
/protein\_id="BAB04000.1"  
/db\_xref="GI:10172894"  
/translation="MAAIVELTDSKSFHSEYVLRDPNLTIDGEMAITTGSQSK  
TTLINIGLEKEDRGVTRLFGRKNPAVHSKQALQFRSKTAYLFQNFALMDETVSK  
NLDVLEILNGESRIVYQOKKKEVLEKVLFNKLNKVAHSLSGGQDQVYATARLLRQC  
DILLADEPGLSDMANRNVIIIDLLCOLNKEGMTIVIVTHPEVADRCDRVITL"  
complement(4354. 4641)  
/gene="BH0282"  
complement(4354. 4641)  
/gene="BH0282"  
complement(4354. 4641)  
/gene="BH0282"  
/note="unknown conserved protein in B. subtilis"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB04001.1"  
/db\_xref="GI:10172895"  
/translation="MKIEHVAIWKDEAMKTEFYVYFENGHANSYHNKKKESFYFL  
TFDGARLEIMQSGSIDKRPDQDGMVAHILSLGSKAEAVNMGRLLHNDG"  
complement(4740. 5537)  
/gene="BH0283"  
complement(4740. 5537)  
/gene="BH0283"  
/note="unknown conserved protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB04002.1"  
/db\_xref="GI:10172896"  
/translation="MTTPITYVDATFNQAFKGPAAVCVLPISRDDIMQHVASENN  
LSEVLFHPYDDGSLRFMTPTNTEVDLGGATLASAHILWELDHISAEQPIPTYSG  
LITAKRSEHLPYDESPQKQENYVPEMLIDGIGIOLPYVGRNRPFLLEIDSEGL  
KELANFSELEBIDTRGILIVTSKSTSYEDISCFPPAGVANDPVTGSAHOCLOGY  
WQENKTRKEPLAYQASRKGMLIKLQHDRVFLLGQAVYVLRSTLL"  
6023. 7096  
/gene="BH0284"  
6023. 7096  
/gene="BH0284"  
/codon\_start=1  
/transl\_table=11  
/product="nicotinate-nucleotide-dimethylbenzimidazole  
phosphoribosyltransferase"  
/protein\_id="BAB04003.1"  
/db\_xref="GI:10172897"  
/translation="MOKYKEMIGSINELDRKAQOOMEOHNLTKPLKSGISGLASIAI  
KIATGTGPOEIBDPATYIVWADHGVAAREVSAVSEVYQOAHNFITGGAIAWYS  
RYSNASVOIVDIGVNGTILQDGLINKVRHGTNNMAGQSPMERREALATIALEVEIAVQ  
RAIQNGNADLPIDIVLAKVGLIAGIAGIVLGAANAARIPVLDGFTITVAALTVNR

gene 1335. 3475  
/gene="BH0280"  
/note="unknown"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB03999.1"  
/db\_xref="GI:10172893"  
/translation="MHWSAAAYADFAVITPRIDLLGPDIVLSYLEEATAHQYNILR  
DVQVLGEGKQITHTTFFALLTEKSPLEAIPRHGSRLTSEDTRGDILISRRKKEDEV  
BEIKG1Q1LSQTHTRTPRPLHPQPNVSPYATQVEQAGTVTEQESATLADLFENY  
LBDYVEQLPLADFAVOTEPESHVSTFEDPLPRKTLAMTILGSGTGMXYLIMRKNK  
TVAIYSMHSEFSFKTMRRLIAGCGVAAVYLLPLLPAFLPMQORRYFLPAAVYTS  
WPLLILASIFLFVYSSRSSLSIESIKNNYTKFLFLYSQAKYVALVSYVFLQTA  
DKULTMEERNMTDMAFGKDKGYEPHISYIDDEENMQESGLMOTDISYESLYPAL  
NNAAGLYIDAVLYEEEMLMATVEYSQEHRSIRVNPNYLQAPYVDENNRKIALOEER  
TVNRVLVEYQKDKQEDILAEFKEDYAAWAMPEQVQESYSHLOEQFDTILMASQEE  
YVSFNLYPEQNGNLIDPPILOVLENNSPYMERPFPGSVGNPLPKVILDMWAST  
VSHYILPLRDLQDMDNAHLYTVNQAKDKDSEIQRALTDITTFVLTATLAFMIYQ  
VSHLEPQAKKFKFLRLRFGHSLFRATRYLFWTLATWVILGIALYRHSQTYLAIY  
VVLFLTEVYVSVLMRLQKNKYSVLKGE"

CDS 3479. 4120  
/gene="BH0281"  
3479. 4120  
/gene="BH0281"  
/codon\_start=1  
/transl\_table=11  
/product="ABC transporter (ATP-binding protein)"  
/protein\_id="BAB04000.1"  
/db\_xref="GI:10172894"  
/translation="MAAIVELTDSKSFHSEYVLRDPNLTIDGEMAITTGSQSK  
TTLINIGLEKEDRGVTRLFGRKNPAVHSKQALQFRSKTAYLFQNFALMDETVSK  
NLDVLEILNGESRIVYQOKKKEVLEKVLFNKLNKVAHSLSGGQDQVYATARLLRQC  
DILLADEPGLSDMANRNVIIIDLLCOLNKEGMTIVIVTHPEVADRCDRVITL"  
complement(4354. 4641)  
/gene="BH0282"  
complement(4354. 4641)  
/gene="BH0282"  
complement(4354. 4641)  
/gene="BH0282"  
/note="unknown conserved protein in B. subtilis"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB04001.1"  
/db\_xref="GI:10172895"  
/translation="MKIEHVAIWKDEAMKTEFYVYFENGHANSYHNKKKESFYFL  
TFDGARLEIMQSGSIDKRPDQDGMVAHILSLGSKAEAVNMGRLLHNDG"  
complement(4740. 5537)  
/gene="BH0283"  
complement(4740. 5537)  
/gene="BH0283"  
/note="unknown conserved protein"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAB04002.1"  
/db\_xref="GI:10172896"  
/translation="MTTPITYVDATFNQAFKGPAAVCVLPISRDDIMQHVASENN  
LSEVLFHPYDDGSLRFMTPTNTEVDLGGATLASAHILWELDHISAEQPIPTYSG  
LITAKRSEHLPYDESPQKQENYVPEMLIDGIGIOLPYVGRNRPFLLEIDSEGL  
KELANFSELEBIDTRGILIVTSKSTSYEDISCFPPAGVANDPVTGSAHOCLOGY  
WQENKTRKEPLAYQASRKGMLIKLQHDRVFLLGQAVYVLRSTLL"  
6023. 7096  
/gene="BH0284"  
6023. 7096  
/gene="BH0284"  
/codon\_start=1  
/transl\_table=11  
/product="nicotinate-nucleotide-dimethylbenzimidazole  
phosphoribosyltransferase"  
/protein\_id="BAB04003.1"  
/db\_xref="GI:10172897"  
/translation="MOKYKEMIGSINELDRKAQOOMEOHNLTKPLKSGISGLASIAI  
KIATGTGPOEIBDPATYIVWADHGVAAREVSAVSEVYQOAHNFITGGAIAWYS  
RYSNASVOIVDIGVNGTILQDGLINKVRHGTNNMAGQSPMERREALATIALEVEIAVQ  
RAIQNGNADLPIDIVLAKVGLIAGIAGIVLGAANAARIPVLDGFTITVAALTVNR

[illegible]



Query Match 46.5%; Score 773; DB 1; Length 3128;  
 Best Local Similarity 67.2%; Pred. No. 1,8e-136;  
 Matches 1094; Conservative 0; Mismatches 535; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAAATCAAAATTTTCAGACAGATGCGGTGCTCCATCGTGGCGGAGTTGAT 74  
 DB 1149 ATGGCAAAAGAAATTAATTTCTGGAAGACGACGTGCTCAATGCTGGCGCGGTTGAT 1208  
 QY 75 ATGTGAGATACCGTCAAGAGTAAGCTGTGTCCTAAAGGGGCAATGTTGCTTGA 134  
 DB 1209 CATTGGCAACACAGTCTTAAGACACCTTAAGACCAAGGCGGCAAGCTGTTTGGAC 1268  
 QY 135 AAGCTTTTGTCTCCCTTAATTAATGACGGGGTAACCATTTGCTAAAGATCGAA 194  
 DB 1269 AAGAGTATGGCTCACCGGAAATTAACCAATGACGGTGTCACTATTGGAAAGTCCATTGAC 1328  
 QY 195 TTAGAAGATCATTTTAAACATGGACCAAAATGGTGTCTGAAGTGGCTTCTAAACC 254  
 DB 1329 TTGGAAATCACTACGAAACATGGGCGCCAGTGTGTCGCAAGTGTCTCAAGACA 1388  
 QY 255 AATGATATTGCTGTGATGGGAGCAGTACTGCAACAGTTTGAACAGCATTTTCAT 314  
 DB 1389 AATGACATTGCTGTGATGGTACCAACCGCAGCGTTTGGACAAAGTATTATCCGC 1448  
 QY 315 GAAGACTAAAAAATGTGACAGCAGGTGCTAATCAATTGATTCGTCGAGGCATTGAA 374  
 DB 1449 GAAGGTATGAAGACGTTACAGCCGGTGGCTAATCTGTTGGCATTTGGCAAGGATTTGA 1508  
 QY 375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGATCTGGCAG 434  
 DB 1509 AAGCACTAAGGCTCCGCTTGACGAATTGCAACAGATTAGCCACCAAGTTAATGGTAAG 1568  
 QY 435 GAAGTATTGCTGATCGCTGCGATCATCATCGCTCGAAGAAATGGTGAAGATATAC 494  
 DB 1569 AAGAAATGCGCAGGTGTGCTGCTTCTCTCAAAATACAGAGTGTGATCTGAT 1628  
 QY 495 TCAGAGCTATGAGAGCGTGTGGGCAACGATGTGATTAACATGCAAGATTCGAGAT 554  
 DB 1629 GGTGACGCCATGSAAGAAAGTTGGCCACGATGGTGAATTACATTGAAGAAACCAAGG 1688  
 QY 555 ATGGAACAGAACTGAAGTGTGAAGCAGTCAATTTGACCGGTGTTACCTGCTCAA 614  
 DB 1689 ATTGACACTGAACCTCTCTGTTGTAAGGAGATGACGTTGATGCGGCTATCTGACCCA 1748  
 QY 615 TACATGTCACAGACATGAAAAAATGGTTGACAGCTGGAAGAAACCATTTATCTAATC 674  
 DB 1749 TACATGTTACCGATATATATATAGATGAAGCTGACCTTGACATCTTATATCTTGATC 1808  
 QY 675 ACGGATAAAAAAGTGTCAAAATCCAGACATTTTGCACACTACTTGAAGAAAGTTCTTAA 734  
 DB 1809 ACCGACAAAAAAGATTTCCAAATTTTCAGGACATTTTGCCTGTTTACAGAAATCGTTCA 1868  
 QY 735 ACCAACCGTCACTACTCTATTTATGCAATGATGTGATGGTGAACACTTCCAAACCTT 794  
 DB 1869 CAAGGTAAAGCAGCTGTGATCACTTGTGACGAGCTGTGTAACCATTTGCCAACCTTA 1928  
 QY 795 GTCTTAAACAGATGCTGCTACTTTCATGTTGCTGCTGTCAAACGCGAGGATTTGG 854  
 DB 1929 GTTCTGAACAAAGTTGCTGTACTTCAATGTTGTGCGGTTAAGGCCCTTGCTTTGG 1988  
 QY 855 GATCGTGTAAAGCTATGCTTGAAGCATTTGCTATCTTGAAGGTGTAACAGTATTA 914  
 DB 1989 GATCGGCTGAAGGCTCACTTGAAGATTTTGTACTTGTGCTGTGTTACCGGTGATCA 2048  
 QY 915 GAGGATCTAGGACTGATTAAGAAATGCTACAGTACAGCCCTTGGACAGGCTGTAA 974  
 DB 2049 TCTGACCTTGGCTGATCTTAAGACACCAAGCTTGAACAAATTTGGTGGCGCGCAAG 2108  
 QY 975 ATTACAGTGAATAAGATGACAGATATTTGTAAGGTTCAGGAGTTCAGAGCTATT 1034  
 DB 2109 GTAAACGTTTACCAAGGACAAACACAGATTGTTGATGGCGCTGCTTAAGCATGCCATT 2168

QY 1035 GCTAACCGTATTCACATGATTAAATGCAATWTAGAAACAACTTCTGACTTTGACCGT 1094  
 DB 2169 GCTGAACGGTTAATATCATCAAGAAAGCAATGTGATGACGACGAGTACTTCGACCGT 2228  
 QY 1095 GAAAACTACAGAACGTTTGGCGAAATTAAGTGTGTGTAGCTGTATCAAGTAGGA 1154  
 DB 2229 GAAAGTTGCAAGAGCGTTTGGCTAAGTTAGCTGGGCGGTGTCCGTTGCAAGTTGGT 2288  
 QY 1155 GCTCCAAAGAGACGCTTAAAGAAATGAACCTGCGCTTGAAGATGCTAATATGCT 1214  
 DB 2289 GCGGCTACTGAAGATTAATGAAGAAACGCAATATGCGATGCAAGATGCTTTGAATGCG 2348  
 QY 1215 ACACGTGACACCGCTTGAAGAGATATGCTGTGCTGTGAGACACACACTATTACGTT 1274  
 DB 2349 ACCGTGCGCGGTTGAAAGAGTTAGCTTGTGCGGCGGCTACAGCTCGTGTGATGTT 2408  
 QY 1275 ATTGAAAAAGTACGACCTCTTGAGCTTTGAGGGCGATGATCTACTGACGTAACATTGTG 1334  
 DB 2409 TTGCTGCTGAGTGTCTGCTTTGAAAGAAAGAGCGACGCTTCAAAACCGGATCAACATCTG 2468  
 QY 1335 CTTCGCTGCTAGAGAGCGCTGTACGTCAAATGCTTTAAATGCTGGTACGAAGCTCC 1394  
 DB 2469 CTACGTGCGCTTGAAGACCTGTTCCGAGATTGCTGAGACCGCGGCAAGAGGTTCT 2528  
 QY 1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGACAGAAACAGATTTAATGCTGCAACAGT 1454  
 DB 2529 GTCATGCTTGAACAGCTGAAGAAAGCAAGGAGGCTGTTCATGATGCGCTACTGAT 2588  
 QY 1455 GAGTGGCTGTATGATTTAAACAGATATTTGACCCCTGTCAAGTAAACAGATACCG 1514  
 DB 2589 GAATGGAGATATGCTTAAGCTAAGCTATTTATGACCCCAACCAAGTGAACCCCTTCTGCA 2648  
 QY 1515 CTTCAAATGACACTTCTGATGATCTATTTTGAACAACAGACAGTGTGCTAAT 1574  
 DB 2649 TTGCAAAAGCCGCTTCCGTTGCTGCTGTGATGCTACAGCTGAAGCCGTTGCTGAT 2708  
 QY 1575 AAACCTGAACAGCTACGCGCAGCGCCAGACATGCCAGAGTATGATCCAGGAATGATG 1634  
 DB 2709 AAGCCAGATCCAAATTAATATATATATGCTGCGCGTGTGCTTAAACCCAGAGGATG 2768  
 QY 1635 GGTGGGATG 1643  
 DB 2769 GCGGCTATG 2777

RESULT 14  
 AF214488  
 LOCUS AF214488 2753 bp DNA BCT 13-JUN-2000  
 DEFINITION Lactobacillus johnsonii groESL operon, complete sequence; and  
 unknown gene.  
 ACCESSION AF214488  
 VERSION AF214488.1 GI:8489166  
 KEYWORDS  
 SOURCE Lactobacillus johnsonii.  
 ORGANISM Lactobacillus johnsonii  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 Lactobacillus.

REFERENCE 1 (bases 1 to 2753)  
 Walker,D.C., Girgis,H.S. and Klaenhammer,T.R.  
 The groESL chaperone operon of Lactobacillus johnsonii  
 JOURNAL Appl. Environ. Microbiol. 65 (7), 3033-3041 (1999)  
 MEDLINE 99318659  
 PUBMED 10388700

REFERENCE 2 (bases 1 to 2753)  
 Walker,D.C., Girgis,H.S. and Klaenhammer,T.R.  
 Direct Submission  
 Submitted (09-DEC-1999) Food Science, North Carolina State  
 University, Room 341 Schaub Hall, Raleigh, NC 27695-7624, USA  
 FEATURES  
 Location/Qualifiers  
 1..2753  
 /organism="Lactobacillus johnsonii"  
 /strain="VPI 11088"

```

CDS      /db_xref="taxon:33959"
          complement(179..331)
          /note="orf"
          /codon_start=1
          /transl_table=11
          /product="unknown"
          /protein_id="AAF5591.1"
          /db_xref="GI:8489167"
          /translation="MTAFGAVLAFISFLDCIVIAFIKLTINDTNKLEAVKRLIETLH
          DLWKKR"
gene      652..2598
          /gene="groESL operon"
gene      652..936
          /gene="groES"
          /gene="groES"
          /gene="groES"
          /note="chaperone protein"
          /codon_start=1
          /transl_table=11
          /product="GroES"
          /protein_id="AAF5592.1"
          /db_xref="GI:8489168"
          /translation="MLQPIGDRVIVVKKDEEEKVGVIVLASNAKEKRPQGEIIVGN
          GKRNSDLIPMSVANKGEIVFDPKYSTNLKTEBKTLVRESDLAVK"
          967..2598
          /gene="groEL"
          /gene="groEL"
          /gene="groEL"
          /note="chaperone protein"
          /codon_start=1
          /transl_table=11
          /product="GroEL"
          /protein_id="AAF5593.1"
          /db_xref="GI:8489169"
          /translation="MAKEIKFSENAHSILKGVKDLADVYKTLIGPKRVNVEKSG
          APTITDGVITAKSIELEHFEEMAKLVSEAOKTNDIGDGTITATVLTQAVREG
          MKNVTAGANPVGIRGLIETVAKAVDELHISKRVSKDIADVASVSASSTEYGNLI
          RDAMEYVGHGVITIEESKIDPELSVEEMGDFRGILSOYMTDNDKMDADNPYI
          LITDKISNIDILPLLOIVQOKSLIILADVDGALPTLVINKRGFENVAVKA
          PGFDRRKMLIEDIAILTGSTVSSDGLKDKPIDGLKAKRVYKSTTVEGA
          GSKFAIERVDQIKKOJADTSPDRKLOERLAKLAGVAIVKGAITTELKERY
          RIDDALNATAAEVGGTALVDVMSIGGVKGDSEDETVKATIMKALGAV
          RQIAENAGKGAVIDLHLEHDEPEVGNATNKNWNNVAKGIIDPTKVTLSALONAS
          IAAILTEVVDADPEDDKQAPAAPNPGMGGM"
BASE COUNT      901 a      458 c      571 g      823 t
ORIGIN
Query Match      46.4%; Score 771; DB 1; Length 2753;
Best local Similarity 68.9%; Pred. No. 4.3e-136;
Matches 1088; Conservative 0; Mismatches 485; Indels 6; Gaps 2;

```

```

Db 1263 TCGTGAAGGTATGAAAGACGTACTGCTGTCGCAAAACCTGTGGTATTCGCGGGTAT 1322
Oy 371 TGAACACGACACAGACAGCGGTGAAGCCCTGGAACCCANTGCTCAACCTGTATCGG 430
Db 1323 TGAACCTGCTACTAAGCAGCGTGTGACGAATTCACAGAAATTAAGCAAGTAAGTAC 1382
Oy 431 CAAGAACCTATTGCTCAGTGCCTGCAGTATCATCAGCTCTGAAAGAGTTGAGAGTA 490
Db 1383 TAAGSAGCAAAATTCGTCAGAGTTCCTCGCTTTCATCAGCTTCACTGAAAGTTGTA 1442
Oy 491 TATCTCAGAACCTATGAGCGCTGGGCAACGANTGTTGATTTACCATCGAAGAAATCTG 550
Db 1443 AATCGGTACCCCAATGGAAGAAAGTTGTCAGCATGGTATTATCTATTCGAAAGTCAA 1502
Oy 551 AGCTATGGAACACAGACCTGTAAGCTGTTGAAGGCATGCAATTTGACCGCTGTACTCTGC 610
Db 1503 GGGTATTGATCTGAACTTGAAGTGAAGTAAGTAAGTAAGTTCGATCGGTACTATTC 1562
Oy 611 TCAATACATGGTCAACAGCAATGAAAGAAATGTTGCAGACCTGGAAGCCATTATCTT 670
Db 1563 ACAATACATGTAAGTCAACAGCAATGGAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1622
Oy 671 AATCAGGATTAAGAAAGTGTCAAAATCAGACATTTTGGCAGCTACTTGAAGAAATCT 730
Db 1623 GATTACTGACAAAGAAAGTTCATTAATATTCAGAAATCTTGCATATTAACAAGAAATCG 1682
Oy 731 TAAACCAACCGCTCATTTACATTTATTCAGATGATGATGATGATGATGATGATGAT 790
Db 1683 TCAACAGGATTAAGAAAGTTCATTAATATTCAGAAATCTTGAAGTTCAGAAATCT 1742
Oy 791 CCTGTCTTGAACAAAGATTCGCTGATTCATTAATGATGATGATGATGATGATGATGAT 850
Db 1743 TCTTGTCTTGAACAAAGATTCGCTGATTCATTAATGATGATGATGATGATGATGAT 1802
Oy 851 TGTGTGATGCTGAAGATTCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 910
Db 1803 TGTGTGATGCTGAAGATTCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1862
Oy 911 TACAGAGATCTAGAGCTTGAATTAAGAAATGCTCAATGATGATGATGATGATGATGAT 970
Db 1863 TTTCTCAACACTTGAATTCGCTGATTAAGAAATGCTCAATGATGATGATGATGATGAT 1922
Oy 971 TACGATTCAGTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
Db 1923 CAGGTTACTGTAACCAAGGATTCGCTGATTAAGATGATGATGATGATGATGATGAT 1982
Oy 1031 TATTGCTAACCGTATTCGATGATTAATTCGCAATTAAGAAACAACTTCGATGATGAT 1090
Db 1983 TATTGCTAACCGTATTCGATGATTAATTCGCAATTAAGAAACAACTTCGATGATGAT 2042
Oy 1091 CCGTGAAAACTACAAGAGCTTTGGCGAAATTTAGCTGGTGTGATGATGATGATGAT 1150
Db 2043 CCGTGAAAACTACAAGAGCTTTGGCGAAATTTAGCTGGTGTGATGATGATGATGAT 2102
Oy 1151 AGGAGCTCAACAGAGAGCTTAAAGAAATGAAATTCGCAATTCGCAATTCGCAATTC 1210
Db 2103 TGTGTGCTGCTGTAACCTGATTAAGAAAGAAAGTACAGAAATCGAAAGCGCTTGA 2162
Oy 1211 TGTACACGTCAGCGCTTGAAGAGGATGATGATGATGATGATGATGATGATGATGAT 1270
Db 2163 CGCAACCGCTGCTGCTGTAAGAAAGTATGATGATGATGATGATGATGATGATGAT 2222
Oy 1271 GGTATTTGA---AAAGTAGAGCTTTGAGCTTGAGCGCATGATG---TACTGACG 1324
Db 2223 TGTATGAAAGTCCATCCAGTACTGTTAAGGATGACAGCAAGCAAGAACTGGT 2282
Oy 1325 TACATGCTGCTGCTGCTGTAAGAAAGCTGTCAGTAATTCGTTAATTCGCTGCTG 1384
Db 2283 TACATGCTGCTGCTGCTGTAAGAAAGCTGTCAGTAATTCGTTAATTCGCTGCTG 2342
Oy 1385 CGAAGGCTCCGTAATTTATGCAAGTTGAAAGAAAGAGCCCTCAGAGAACAGGATTTAAT 1444

```

Db 2343 AGATGCTGCTTATCTTAGACACTTAGACATGAAGACCAGAGTTGGTTACAATGC 2402

QY 1445 TGCACAGCTGAGTGGTGGATGATGATTAACAAGACATCTTACCCTGTCAGTAAC 1504

Db 2403 TGCACACTAACAGTGGGAAACATGTTAAGCTGGATATATGACCACCACTAAGGTAAC 2462

QY 1505 AGCATAGAGCTTCACAAATGACACTCTCTAGTACTGCTTATTTTGACACAGAGACT 1564

Db 2463 TCGTTCAGCATTCAAAATGCTGCTTCAATTCGCTGCTGTTATTAACACTAGAGCTGT 2522

QY 1565 TGTTCCTAATAACCTGAA 1583

Db 2523 TGTTCCTGACGCTCCAGAA 2541

RESULT 15  
AF269843/C 3885 bp DNA BCT 01-AUG-2000  
LOCUS Staphylococcus epidermidis strain SRI clone step.1026c04 genomic  
DEFINITION sequence.  
ACCESSION AF269843  
VERSION AF269843.1 GI:9623742  
KEYWORDS  
SOURCE Staphylococcus epidermidis.  
ORGANISM Staphylococcus epidermidis.  
REFERENCE 1 (bases 1 to 3885)  
AUTHORS Kimmerly,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M.,  
Rubino,M.A., Nelson,F.J., Rivers,P.R., Tortuella-Miller,I.,  
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,  
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and  
Furdon,P.J.  
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis  
genome

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3885)  
AUTHORS Taylor,J.David, Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,  
Rubino,M.A., Nelson,F.J., Rivers,P.R., Tortuella-Miller,I.,  
Listenebee,S., Ashanti,C., Altschuller,G., Mammo,L., Shepherd,N.S.,  
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and  
Furdon,P.J.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and  
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore  
Drive, Research Triangle Park, North Carolina 27709-3398, USA  
FEATURES  
source 1..3885  
/organism="Staphylococcus epidermidis"  
/strain="SRI"  
/db\_xref="taxon:1282"  
/clone="step.1026c04"  
BASE COUNT 1061 a 739 c 560 g 1525 t  
ORIGIN

Query Match 45.5%; Score 756.4; DB 1; Length 3885;  
Best Local Similarity 67.5%; Pred. No. 2.3e-133;  
Matches 1063; Conservative 0; Mismatches 511; Indels 0; Gaps 0;

Db 2691 GTTAGAAGATCATATGAGATATATGGGTGCAAAATTAGTGCAGAACTTCGANTAAAC 2632

QY 254 CAATGATATTTGCTGATGATGAGACACTACTGCAACAGCTTTTGACACAAACCATTTGTCA 313

Db 2631 AATGGAATGCTGGGAGCGGTACMACTACAGACAGCTTTAGCACATCAATGATTC 2572

QY 314 TGAAGACTAAATAATGTGACAGCAGGTGCTAATCCAAATTTGATTCGTCGAGGACTGA 373

Db 2571 GGAAGGCTTAAAGAAATGTTCAAGTGGTCAAAATCCGTGAGGCTTAAGACAGATTTGA 2512

QY 374 AACAGACAGCAACAGCTGTGTAAGCTTGAAGGCATTCGTCMACTATCTGGCA 433

Db 2511 CAAGAGTGCAAGTGGCTATAGAGGCTTCAGATTTCTCAAAAGGTGAATAA 2452

QY 434 GGAAGCTATTTGCTCAGGTCGCTCAGATCATACCGCTGAAAAAGTTGAGATATAT 493

Db 2451 GAACGAGATGCGCCAGAGTTGAGCTATTTAGACAGCAGATGAAGAAATGCTGCTACAT 2392

QY 494 CTCAGAGCTATGAGACGCTGTGGCAACGATGCTGATTAACATCGAAGATCTGAGG 553

Db 2391 TTTCTGAAGCAATGATTAAGATAGTAAACGATGCGCTTATCACTATTAAGAAATCAATG 2332

QY 554 TATGGAACGAACTTGAAGTGGTTGAAGCATGCAATTTGACCGTGTACTGTCTCA 613

Db 2331 GTTAAATACGAATTAAGATAGTTGAAGAAATGCAATTTGATCGGCTATCAATCACC 2272

QY 614 ATACATGCTCACACAAATGAAGAAATGCTTGACAGCTTGAAGAACCATTTATCTTAAT 673

Db 2271 ATATATGATAGTACTGACTGACATGATTAATGATAGCTGATTTGAACGTCATATATTAAGT 2212

QY 674 CACGATTAAGAAAGTGTCAAAATCCAGACATTTTGCACACTTGTAGAGAGTCTTAA 733

Db 2211 AACGATTAAGAAATTTTCAATTCATTCACAGATATTTCCATTTATTAAGAACAGTTGCA 2152

QY 734 AACCAACGCTTCATTTACTATTTTGCAGATGATGATGATGATGATGATGATGATGATGAT 793

Db 2151 GGTATGTCGACCAATTTTATTTGTTGCGATGAAGTGAAGGAGTACATTTCTATAT 2092

QY 794 TGTCTTGAACAAAGTCTGCTGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853

Db 2091 TGTCTTGAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032

QY 854 TGAATGCTGTAAGTATGCTGTAAGACATTTGCTATCTTGCAGAGTGTGATGATGATGAT 913

Db 2031 TGAATGCTGTAAGTATGCTGTAAGACATTTGCTATCTTGCAGAGTGTGATGATGATGAT 1972

QY 914 AGAGATCTAGAGCTGATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 973

Db 1971 TGAATGCTGTAAGTATGCTGTAAGACATTTGCTATCTTGCAGAGTGTGATGATGATGAT 1912

QY 974 GATTAAGTGTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033

Db 1911 AGTTGAAGTGTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1852

QY 1034 TGTCAACGCTATGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093

Db 1851 TGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1792

QY 1094 TGAATACTACAGACGCTTTGGCGAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153

Db 1791 AGAAATACTACAGACGCTTTGGCGAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1732

QY 1154 AGCTCCACAGACAGCTTTTAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 1213

Db 1731 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1672

QY 1214 TACAGCTGACGCTTTGAAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273

Db 1671 AACAGCTGCGCGGCTGGAAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612

QY 1274 TATGAAAAGTACAGCTCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1333



```

Db 1611 ATATCAAAAAGTAAGTGAATTAAGCAGAGGATGATGTTGAAAGGGGTGTTAATATCGT 1552
QY 1334 GCTTCGTCCTAGAAAGCGCTGACGTCAAAATGCTTTAAATGCTGGTAGAAGGCTC 1393
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1551 ATTAAAGCATTAACAGCAGCTGTGACAAATTGCTGAAATGCAAGATTAGAGGCTC 1492
QY 1394 CGTAGTTATTGACAGTTGAAAAACAGCCCTGCAGAGACAGATTTAATGCTGCACAGG 1453
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1491 AATTATGTTGAACGTTTAAACATGCTGAAGCGGGCGGTGGTTCAATGCAGCACAAA 1432
QY 1454 TGAAGTGGGTGATATGATTAACAGAAATCATGACCTGTCAAAATGTAACAGATCAGC 1513
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1431 TGAATGGGTTAATATGTTAGAGAGGTAATGATGATCAACTAAAGTAACCTCGTTCAGC 1372
QY 1514 GCTTCAAAATGACAGCTCTGTAGCTAGCTTATTTTGACAAGACAGAGCTGTGCTAA 1573
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1371 GTTACAAACATGCAGCAGGTAGCTGCTATGTTCTTAACAACAGCAAGCAGTGTGCTAG 1312
QY 1574 TAAACCTGAACCAAG 1587
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1311 TATTCAGAGCCAG 1298

```

Search completed: February 15, 2001, 23:20:43  
 Job time: 8971 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 19:54:40 ; Search time 97.06 Seconds  
(without alignments)  
6428.763 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661  
Sequence: 1 GAATTCGGCTTCATATGCA.....TGGCGGATTAACCGCAATTC 1661

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_36.\*  
2: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /cgn2\_2/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /cgn2\_2/gcgdata/geneseq/geneseqn/NA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1661	100.0	1661	20	X86155
2	1052	63.3	1654	20	X86153
3	1035.4	62.3	5365	19	V52210
4	942.6	56.7	3625	20	X12979
5	704.2	42.4	1647	18	T69201
6	704.2	42.4	4972	18	T69203
7	590.6	35.6	1665	20	X86152
8	568.2	34.2	1662	20	X86154
9	552.8	33.3	1635	21	X254509
10	551.6	33.2	910715	20	X20248
11	550	33.1	2465	12	Q13136
12	536.8	32.3	1635	21	A13000

13	536.8	32.3	1635	21	254508	Neisseria meningit
14	528.6	31.8	1838	14	Q47926	hsp60 DNA, Helico
15	528.6	31.8	1838	14	Q48734	Hsp gene, Helicob
16	528.6	31.7	2223	12	O13137	Hyp operon, Chlam
17	527	31.7	1724	19	X14495	H. pylori GHPO 118
18	526.6	31.7	1635	21	254507	Neisseria gonorrhe
19	505.8	30.5	1038602	20	Z01425	Complete genome se
20	496.6	29.9	2284	16	Q90181	Helicobacter pylori
21	496.6	29.9	2284	17	T45681	H. pylori heat sho
22	496.6	29.9	2322	16	O75321	Heat shock protein
23	483.2	29.1	3613	9	N80339	Clone Y3178 Insert
24	470	28.3	1017	18	V74577	Staphylococcus aur
25	458.6	27.6	4260	9	N81768	Sequence encoding
26	458.6	27.6	4380	9	N80222	Sequence of Mycob
27	458.6	27.6	4380	19	V05708	Mycobacterium tube
28	446.8	26.8	1626	20	Z11371	Nucleotide sequenc
29	445.8	26.8	1647	17	T14265	Brevibacterium fla
30	444.8	26.8	1569	19	V34608	M. vaccae antigen
31	444.8	26.8	1569	20	Z11343	Nucleotide sequenc
32	441.4	26.6	580073	18	T58840	Mycoplasma genital
33	432.6	26.0	2193	11	O04669	Heat shock protein
34	420.6	25.3	2668	13	O22485	groEL-1 gene. Str
35	419.8	25.3	1620	13	O22482	groEL-1 gene codin
36	409.6	24.7	2242	18	T58403	Human heat shock p
37	362.8	21.8	2167	13	O22484	groES el operon w
38	354	21.3	1320	13	O22481	Heat shock protein
39	315.8	19.0	1777	19	V59425	Heat shock protein
40	296.4	17.8	831	20	V90856	Nucleotide sequenc
41	273.8	16.5	985	20	Z11372	Nucleotide sequenc
42	271.8	16.4	927	19	V34610	M. vaccae antigen
43	271.4	16.4	927	20	Z11345	Nucleotide sequenc
44	217.4	13.1	692	20	X20722	Polynucleotide seq
45	198.8	12.0	544	18	V75108	Staphylococcus aur

## ALIGNMENTS

RESULT	1
ID	X86155
XX	X86155 standard; DNA: 1661 BP.
XX	
AC	X86155;
XX	
DT	22-SEP-1999 (first entry)
XX	
DE	DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.
XX	
KW	Heat shock protein; Hsp60-2; immune response; immunological carrier
KW	cancer control; tumour; sarcoma; cancer; gene therapy; ss.
XX	
OS	Streptococcus pyogenes.
XX	
PN	WO9935270-A1.
XX	
PD	15-JUL-1999.
XX	
PE	29-DEC-1998; 98MO-CA01203.
XX	
PR	31-DEC-1997; 97US-0001737.
XX	
PA	(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX	
PI	Mizzen L, Wisniewski J;
XX	
DR	WPI; 1999-430397/36.
DR	P-PSDB; Y23904.
XX	
PT	New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT	useful in vaccines, as carriers for other immunogens, as anticancer
PT	agents and for diagnosis
XX	
PS	Claim 3; Fig 4A-B; 176pp; English.

XX The present sequence encodes a heat shock protein, designated Hsp60-2.  
 CC The protein, its fragments, variants and fusion proteins, are  
 CC used to elicit or enhance an immune response against Streptococcus,  
 CC and to elicit a similar response to a target antigen fused to the  
 CC protein. Unlike other immunological carriers, Hsp60 proteins are not  
 CC immunosuppressive so provide an increased response to any conjugated or  
 CC fused antigen. Also, where used for cancer control, they lack the side  
 CC effects associated with endotoxins. They can also be used to detect  
 CC specific antibodies and in treatment or prevention of tumours  
 CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or  
 CC liver). The Hsp60 polynucleotide is used for recombinant production  
 CC of the protein, as a source of primers and probes for detecting  
 CC streptococci in standard hybridization/amplification assays, and  
 CC therapeutically in gene therapy vectors.

SQ Sequence 1661 BP; 512 A; 308 C; 407 G; 434 T; 0 other;

Query Match 100.0%; Score 1661; DB 20; Length 1661;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCTTCATATGGCAAAAGAAATCAATTTTCAGAGAGATGGCGCTGCGCCATGG 60  
 DB 1 gaattcggcttcataatggcaaaaagaaatcaattttcagagagatggcgctgccccatgg 60  
 QY 61 TGGCGGAGTGTATGATAGAGATACCGTCAAGTAACGCTTGCTCAAAAGGGCGCA 120  
 DB 61 tggcggagtgtatgattagagataccgtcaagtaacgcttggctcaaaaggcgca 120  
 QY 121 ATGTTTCTTCTGAAAAAGCTTTTGTTGCTCCCTTAATTAATGACGGGGTACCATTTG 180  
 DB 121 atgtttcttctgaaaaagcttttggttgcctcttaatttaattgacggggtaccat 180  
 QY 181 CTAAAGAGATCGAATTAAGATCATTTTGAAACATGGAGCAAAATGCTGTCGAAG 240  
 DB 181 ctaaagagatcgaaatgaagatcattttgaaacatggagcaaaatgctgtcgaag 240  
 QY 241 TGGCTTCTAAACCAATGATATTGCTGTGATGGAGGAGTACTGCAACAGTTTGAAC 300  
 DB 241 tggcttctaaaccaatgatattgctgtgatggaggtactgcaaacagtttgaac 300  
 QY 301 AAGCCATTGTTCTGTAAGACACTAAAAATGTCACAGCGTGTCAATTCATTTGATTC 360  
 DB 301 aagccattgttctgtaagacactaaaaatgttcacagcggtgtcaatttcattg 360  
 QY 361 GTCGAGGCAATTAAGACAGCAACAGCTTTGAAGCTTGAAGCCATTGCTCAAC 420  
 DB 361 gtcgaggcaatttaagacagcaaacagcttgaagccttgaagccattgctcaac 420  
 QY 421 CTGTATCTGGCAAGAACTATTGCTCAGGTGCTGCAGTATCATCAGCTTGAAAAAG 480  
 DB 421 ctgtatctggcaagaactattgctcaggtgctgcagtatcatcagcttgaagaa 480  
 QY 481 TTGGAGAGATATCTAGAAGCTATGAGCGTGTGGCAACGATGGTGTATACATCG 540  
 DB 481 ttggagagatattctagaagctatgagcgtgtggcaacgattggttattacatcg 540  
 QY 541 AAGAACTCTGAGTATGAAACAGACTTGAAGTGTGTAAGCAGTCAATTGACCGTG 600  
 DB 541 aagaaactctgagtatgaaacagacttgaagtgtgtaagcagtcaattgaccgtg 600  
 QY 601 GTTACCTGCTCAATACAGGTGCACAGACATGAAAAATGTTGCGAGCCTTGAAGACC 660  
 DB 601 gttacctgctcaatacaggtgcacagacatgaaaaatgttgcgagccttgaagacc 660  
 QY 661 CATTTATCTTAAATCAGCGATAAAAAGTGTCAAACTCAAGACATTTTGCCACTACTTG 720  
 DB 661 catttatctttaaatacagcgataaaaaagtgtcaaaactcaagacattttgccact 720  
 QY 721 AGGAAGTTCTTAAACCAACCGTCCATTTACTATTTCAGAGATGATGTGATGGTGAAG 780  
 DB 721 aggaagttctttaaaccacacgctccatcttactcatlctgcagatgatgtgagtgaag 780

DB 721 aggaagttctttaaaccacacgctccatcttactcatlctgcagatgatgtgagtgaag 780  
 QY 781 CACTTCCAACCCCTTGTCTTGAAGAGATTCGTGATCTTCAATGTGTGCTGTCAAG 840  
 DB 781 cacttccaaacccttgtcttgaagagattcgtgattcttcaatgtgtgctgtcaag 840  
 QY 841 CGCCAGGATTTGGTATCGTCGTAAGCTTATGCTTAAGACATTTGCTTATGACAGTG 900  
 DB 841 cgccaggatttggtatcgtcgttaagcttattgctttaagacatttgcttattgac 900  
 QY 901 GTACAGGATTTACAGAGATCTGAGACTTAATTAAGATGCTACAAATGACAGCCCTTG 960  
 DB 901 gtacaggttttacagagatctgagacttaattaaagtctacaaatgacagcccttg 960  
 QY 961 GACAGCGTCTAAGATTACAGTTGATTAAGATAGCACAGTAAATGTTGAAGTTACGAA 1020  
 DB 961 gacagcgctctaagattacagttgatattaagatagcacagttaattgttgaaagtt 1020  
 QY 1021 GTTCAGAGCTATTGCTAACCGTATTGCACTGATTAATGCAATTAAGAAACAACTT 1080  
 DB 1021 gttcagagctattgctaacgctattgcaactgatttaattgcaatttaagaaacct 1080  
 QY 1081 CTGACTTTCGCGTGAAAAAACTAAGAAAGCTTTGGGAAATTAAGCTGGTGTAGCTG 1140  
 DB 1081 ctgacttttcgctgtaaaaaactaagaaagctttgggaaatthaagctgggtgtag 1140  
 QY 1141 TTATCAAGATGAGAGCTCCAAACAGACAGACTTAAAGAAAGAACTTCGATTGAG 1200  
 DB 1141 ttatcaagatgagagctccaaacagacagactttaaagaaagaaacttcgattgag 1200  
 QY 1201 ATGCTCTAAATGCTACACGTGACAGCCGTTGAGAAGCTATGTTGCTGTGTGAACAC 1260  
 DB 1201 atgctctaaatgctacacgtgacagccgttgagaagctatgttgcgtgtgaacac 1260  
 QY 1261 CACTTATTCAGTATTAAGAAAGTGAAGCAAGCTTACAGCTTGAGGGGAGATGATGAC 1320  
 DB 1261 cacttatctcagatattaagaaagtgaagcacttcagcttgaggcgatgacgactg 1320  
 QY 1321 GACGTAACTATTGCTGCTGCTCTAGAAAGAGCCTGTACGTCAAAATGCTTAAATGCTG 1380  
 DB 1321 gacgtaaactattgctgctgctctagaaagagcctgtacgtcaaaatgctttaa 1380  
 QY 1381 GGTACAGAGGCTCCGTATGATTAAGCAAGTTGAAAAACAGCCCTGCAGACAGATTTA 1440  
 DB 1381 ggtacagaggctccgtattgattaaagcaagttgaaaaacagccctgcagacag 1440  
 QY 1441 ATGCTCAACAGAGTGAAGTGTGATGATTAAGCAAGATCATTTGACCCGTCAAG 1500  
 DB 1441 atgctcaacagagtgaagtgtgattaaagcaagatcatttgacccgtcaag 1500  
 QY 1501 TAACAGATCAGCGCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTTGACAACGAA 1560  
 DB 1501 taacagatcagcgcttcaaaatgacagcttctgtagctagcttattttgacaac 1560  
 QY 1561 CAGTTTGTCTAATAACTGAACCTGAACGCTACGACCGCAGCAATGCCAGAGTATG 1620  
 DB 1561 cagtttgtcttaataactgaacctgaacctgacgacgcaaatgccagagtattg 1620  
 QY 1621 ATCCAGAAATGATGGTGGATGGGCGGATTAAGCCCAATTG 1661  
 DB 1621 atccagaaatgatgggtggatgggcggttaagcccaattg 1661

RESULT 2  
 X86153  
 ID X86153 standard; DNA; 1654 BP.

XX X86153;  
 AC 22-SEP-1999 (first entry)  
 DT  
 XX  
 DE DNA encoding a Streptococcus pneumoniae heat shock protein (Hsp)60-2.  
 XX

KW Heat shock protein: Hsp60-2; immune response; immunological carrier;  
 KM cancer control; tumour; sarcoma; cancer; gene therapy; ss.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN W09935270-A1.  
 XX  
 PD 15-JUL-1999.  
 XX  
 PF 29-DEC-1998; 98MO-CA01203.  
 XX  
 PR 31-DEC-1997; 97US-0001737.  
 XX  
 PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.  
 XX  
 PI Mizzen L, Wisniewski J;  
 XX  
 DR WPI: 1999-430397/36.  
 DR P-PSDB: Y23902.  
 XX  
 PT New nucleic acid encoding heat shock protein-60 from Streptococcus,  
 PT useful in vaccines, as carriers for other immunogens, as anticancer  
 PT agents and for diagnosis  
 XX  
 PS Claim 3: Fig 2A-B; 176pp; English.  
 XX  
 CC The present sequence encodes a heat shock protein, designated Hsp60-2.  
 CC The protein, its fragments, variants and fusion proteins, are  
 CC used to elicit or enhance an immune response against Streptococcus,  
 CC and to elicit a similar response to a target antigen fused to the  
 CC protein. Unlike other immunological carriers, Hsp60 proteins are not  
 CC immunosuppressive so provide an increased response to any conjugated or  
 CC fused antigen. Also, where used for cancer control, they lack the side  
 CC effects associated with endotoxins. They can also be used to detect  
 CC specific antibodies and in treatment or prevention of tumours  
 CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or  
 CC liver). The Hsp60 polynucleotide is used for recombinant production  
 CC of the protein, as a source of primers and probes for detecting  
 CC streptococci in standard hybridization/amplification assays, and  
 CC therapeutically in gene therapy vectors.  
 XX  
 SO Sequence 1654 BP; 500 A; 326 C; 401 G; 427 T; 0 other;

Query Match 63.3%; Score 1052; DB 20; Length 1654;  
 Best Local Similarity 77.7%; Pred. No.-8e-266;  
 Matches 1271; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

QY 1 GAATTCGGCTTCATATGAGCAAAATCAATTTTCAGCAGATGCGCTGCTGCATATG 60  
 DB 1 gaattcggcttcataatgcaaaagaattcaatttcacagatgcccgttcagctatg 60  
 QY 61 TGCCGGAGTTGATATGTTGAGATACCGTCAAGTAACCTGTGTCCTAAAGGCGCA 120  
 DB 61 tccggagttgatattgttgagatccgtcaagtaacctgtgctcctaaaggcgca 120  
 QY 121 ATGTTGTTCTTGAAGAAAGCTTTTGGTCTCCCTTAATTAATGACGGGGTAAACATTTG 180  
 DB 121 atgttgcttcttgaagaaagcttttggctcccttaatttaattgacggggtaaacattt 180  
 QY 121 atgctcgtcttgaagaaagcttgcacacctgattaccacaaagcagtggtgactatg 180  
 QY 181 CTAAAGAGATCGAATTAAGATCATTTTGAAGAAATCGAGGAGCAAAATTTGGTGTCTGCAAG 240  
 DB 181 ctaagagatcgaaatgaagatcatcttgaagaaatcgagggcaaaatttgltacagaag 240  
 QY 241 TGGGTCTTAAACCAATGATTTTGTGCTGATGGAGACACTACTGCAACAGTTTGTGACAC 300  
 DB 241 tgggtctttaaaccatgatttttgtgctgattggagacactactgcaacagtttgtgacac 300  
 QY 301 AAGCATTTGTCATGAGAGCATTAATAAATGTGACAGAGGTGCTAATTCATTTGTTATCC 360  
 DB 301 aagcatTTGTCATGAGAGCATTAATAAATGTGACAGAGGTGCTAATTCATTTGTTATCC 360  
 QY 361 GTGAGGCAATTGAAACAGCAACAGCAACAGCTGTTGAAGCCCTTGAAGCCATTTGCTCAAC 420

DB 361 gtgctggatggaacacgcaatgctccgacagcagttgaagcttgaagaaacacgtcatcc 420  
 QY 421 CTGTATCTGCGCAGAGAGCTATTGTCTGCGGTGCGAGTATTCACGCGCTGAAAAAG 480  
 DB 421 ctgtatctgCGCAGAGAGCTATTGTCTGCGGTGCGAGTATTCACGCGCTGAAAAAG 480  
 QY 481 TTGAGAGATATATCTCAGAACCTATGAGCGCTGTGGCGAACGATGTGTGATTACATCG 540  
 DB 481 ttgagagatatatctcagaaacctatgagcgctgtggcgaaCGATGTGTGATTACATCG 540  
 QY 541 AAGATCTCGAGGATATGGAACACAGACTTGAAGTGTGAAGGCATGCAATTTGACCGTG 600  
 DB 541 aagatctcGAGGATATGGAACACAGACTTGAAGTGTGAAGGCATGCAATTTGACCGTG 600  
 QY 601 GTTACCTGCTCAATPACATGTGTCACAGCAATGAAAAATGGTTGACGACTTGAAGAAC 660  
 DB 601 gttacctgCTCAATPACATGTGTCACAGCAATGAAAAATGGTTGACGACTTGAAGAAC 660  
 QY 661 CATTTATCTTAATCAGCATTAATAAAGCTGTCAACATCCAGACATTTTGGCACTACTTG 720  
 DB 661 catTTATCTTAATCAGCATTAATAAAGCTGTCAACATCCAGACATTTTGGCACTACTTG 720  
 QY 721 AGGAAGTTCTTAAACCAACACCGCTCCATTACTCATTTATTTGACAGATGATGTGATG 780  
 DB 721 aggaagTTCTTAAACCAACACCGCTCCATTACTCATTTATTTGACAGATGATGTGATG 780  
 QY 781 CACTTCCAAACCTTGCTGTTGAACAAGATTGGTGACTTCAATAGTGGTCTGCTCAAG 840  
 DB 781 cacttccAAACCTTGCTGTTGAACAAGATTGGTGACTTCAATAGTGGTCTGCTCAAG 840  
 QY 841 CGCCAGAGATTTGATGATGCTGTAAGAGCTATGCTTGAAGACATTTGCTATCTGACAGTG 900  
 DB 841 cgccagagATTTGATGATGCTGTAAGAGCTATGCTTGAAGACATTTGCTATCTGACAGTG 900  
 QY 901 GTACAGTATTCAGAGAGATTCAGACCTTGAATTAAGATGCTACATGACAGCCCTTG 960  
 DB 901 gtaCAGTATTCAGAGAGATTCAGACCTTGAATTAAGATGCTACATGACAGCCCTTG 960  
 QY 961 GACAGGCTGCTAAGATTTACATGTTAAGATPACACAGTAAATGTTGAAGTTGACAGTA 1020  
 DB 961 gacAGGCTGCTAAGATTTACATGTTAAGATPACACAGTAAATGTTGAAGTTGACAGTA 1020  
 QY 1021 GTTCAGAGATTTGCTAATCGATATTTGCATGATTAATTCGCAATTAAGAAACCAACTT 1080  
 DB 1021 gtTCAGAGATTTGCTAATCGATATTTGCATGATTAATTCGCAATTAAGAAACCAACTT 1080  
 QY 1081 CTGACTTTGACCGTGAAGAACTACAGAGACGTTTGGCAATTAAGCTGTGTGATCGTG 1140  
 DB 1081 ctgACTTTGACCGTGAAGAACTACAGAGACGTTTGGCAATTAAGCTGTGTGATCGTG 1140  
 QY 1141 TTTATCAAGTGAAGCTTCCACACAGACAGCTTTAAAGAAATGAAGAACTTGGCATTTGAGG 1200  
 DB 1141 ttTATCAAGTGAAGCTTCCACACAGACAGCTTTAAAGAAATGAAGAACTTGGCATTTGAGG 1200  
 QY 1201 ATGCTTAAATGCTACACAGCTGACCGCTTGAAGAAAGATGATGCTGTGCTGGAACAG 1260  
 DB 1201 atGCTTAAATGCTACACAGCTGACCGCTTGAAGAAAGATGATGCTGTGCTGGAACAG 1260  
 QY 1261 CACTTATTAAGCTTGAAGAAAGTACAGAGCTTTGAGCTTGAAGCGCATGATGCTACTAG 1320  
 DB 1261 cactTATTAAGCTTGAAGAAAGTACAGAGCTTTGAGCTTGAAGCGCATGATGCTACTAG 1320  
 QY 1321 GAGCTAATATGCTGCTGCTCTAGAAAGCCTGTACGTCATTAATTTAATATGCTG 1380  
 DB 1321 gagCTAATATGCTGCTGCTCTAGAAAGCCTGTACGTCATTAATTTAATATGCTG 1380  
 QY 1381 GGTACGAAGGCTCCGATGTTTATGACAAAGTGAAGAAACAGCCCTGACAGAGCAAGATT 1440  
 DB 1381 ggtACGAAGGCTCCGATGTTTATGACAAAGTGAAGAAACAGCCCTGACAGAGCAAGATT 1440  
 QY 1441 ATGCTGCAACAGGTGAGTGGTGTATGATTAATAACAGAAATCATTTGACCCCTGCAAG 1500  
 DB 1441 atGCTGCAACAGGTGAGTGGTGTATGATTAATAACAGAAATCATTTGACCCCTGCAAG 1500

Db 1441 acgcagcaactggcggagtggttaacatgatcatgaagatcatcatgaagcaagtaag 1500  
 QY 1501 TAACAGCATGAGCGCTTCAAAATGCAGCTTCTGTAGCTAGTCTTATTTGACACAGAG 1560  
 Db 1501 tggatcgcttcacccctacaaatgacagcatctgtagccagcttgatttggacaacagaag 1560  
 QY 1561 CAGTTGTGCTAATAACCTGACAGCACTACGCCAGGCCACGACATGCCAGCATGTGC 1620  
 Db 1561 cagtcgtacgcaataaacacagccagtagcccgagctccagcaatgatccaaatga 1620  
 QY 1621 ATCCAGGATGATGG 1636  
 Db 1621 tgggtcggaatcgagcg 1636  
 RESULT 3  
 ID V52210  
 V52210 standard: DNA; 5365 BP.  
 XX V52210;  
 AC V52210;  
 XX 23-OCT-1998 (first entry)  
 DT Streptococcus pneumoniae genome fragment; SEQ ID NO:77.  
 DE Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;  
 XX computer readable medium; vaccine; pharmaceutical composition; ds.  
 KW Streptococcus pneumoniae.  
 OS Streptococcus pneumoniae.  
 XX MO9818931-A2.  
 PN 07-MAY-1998.  
 PD 30-OCT-1997; 97MO-US19588.  
 PF 31-OCT-1996; 96US-0029960.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PA Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;  
 PI Kunesh CA, Rosen CA;  
 PI WPI; 1998-272225/24.  
 DR Computer-readable medium with recorded Streptococcus pneumoniae  
 PT polynucleotide sequences - useful in diagnostic kits and assays, and  
 PT pharmaceutical compositions and vaccines for Streptococcus  
 PT pneumoniae  
 PS Claim 1; Page 628-631; 1409pp; English.  
 XX The present invention describes a computer readable medium which has  
 CC on it, or a representative fragment or a sequence at least 95% identical  
 CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1  
 CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus  
 CC pneumoniae. The present invention also describes an isolated nucleic acid  
 CC molecule encoding a homologue of any of the fragments of the S.pneumoniae  
 CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced  
 CC by a process comprising: (a) screening a genomic DNA library using as a  
 CC probe a target sequence defined by any of the sequences in SEQ ID NO:1  
 CC to 391, identifying members of the library which contain sequences  
 CC that hybridize to the target sequence and isolating the nucleic acid  
 CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced  
 CC from an organism, amplifying nucleic acid molecules whose nucleotide  
 CC sequence is homologous to amplification primers derived from the  
 CC fragment of the S. pneumoniae genome to prime the amplification and  
 CC isolating the amplified sequences. The computer readable medium can be  
 CC used in a computer-based system for identifying fragments of the  
 CC S. pneumoniae genome of commercial importance, or expression modulating  
 CC fragments of the S. pneumoniae genome. Products from the present  
 CC invention can be used in diagnosis kits and assays, and pharmaceutical

CC compositions and vaccines for S. pneumoniae.  
 XX Sequence 5365 BP; 1654 A; 960 C; 1082 G; 1669 T; 0 other;  
 SQ  
 Query Match 62.3%; Score 1035.4; DB 19; Length 5365;  
 Best Local Similarity 77.4%; Pred. No. 2,8e-261;  
 Matches 1255; Conservative 0; Mismatches 366; Indels 0; Gaps 0;  
 QY 14 TATGCAAAAGAAATCAATTTTCAGCATGGCGGTGCGCATGGTGGCGAGTTGA 73  
 Db 277 tatgtcaaaagaataatcaatltcatcagatgccgttcagcaatgcttgcgta 336  
 QY 74 TATGTAGCAGATACCGTCAAGTAACGCTTGTCTTAAGGGCGCAATGTTCTTGA 133  
 Db 337 talcccttgacagacigttaagtaaccttgggaccaaagtcgaatgctgtcttga 396  
 QY 134 AAAAGCTTTTGCTTCCCTTAATTAATTAATGACGGGCTAACCATTTGCTTAAGATCGA 193  
 Db 397 aaagtcattcgttcacaccttgatcccaatgacggtgtgacatttgcgaagaatcga 456  
 QY 194 ATTAGAAGATCATTTGAAACATGGAGCAAAATTTGCTGTGAAGTGCTCTTAAC 253  
 Db 457 attgaaagaccatttgaataatggtctgaagtagtcaagaagttagcttctaacc 516  
 QY 254 CAATGATATTGCTGTGATGAGCAGACTACTGCAACAGTTTGGACAAACCATTTGCTCA 313  
 Db 517 caatgatctcgagtgaggaagtaactgaactgcaacagcttgcacccaagatctgcg 576  
 QY 314 TGAAGGACTAAAAATGTGACAGCAGTGTCTAATTCATTTGATTCGCTGAGCATTGA 373  
 Db 577 tgaaggaatcaaaaacgctcagcaggtgcaaatccaatcgtatcttcgttgcgttga 636  
 QY 374 AACGCAACAGCAAGCTGTGAAGCTTGAAGCCATGCTCAACCTGTACTGTGGCAA 493  
 Db 637 aacagcagttgcgcgacagctgaagcttgaagcttgaataaagccatccgttgcataaa 656  
 QY 434 GGAAGCTATTGCTCAGGTCGCTGAGTATCATCAGCTCTGAAAGATTGGAGATATAT 493  
 Db 697 agaagctatcgtccaagltgcagcgtatcttctgttgcgaagaagtgtgagtaacat 756  
 QY 494 CTCAGAAAGCTATGAGCAGCTGTGGCAACGATGCTGATTAACATGGAAGATCTGAGG 553  
 Db 757 ctctgaagcaatggaataaaglttgcaagaagcgtgtatcatccatcgaagaagtcagctg 816  
 QY 554 TATGAAACAGAACTGAAGTGTGAAGCAGCATGATTTGACCGTGTGATCTGTCTCA 613  
 Db 817 tatggaacagagcttgaagctgtagaagaatgcatgctgttgcagcgtgtgtaaccttaca 876  
 QY 614 ATACATGCTCACAGACAAATGATGTTGCAAGCCTTGAACCCATTTATCTTAAT 673  
 Db 877 gtacatggtgacagatagcgaataaaltggtgctgaccttgaaatccgtacatcttggat 926  
 QY 674 CAGGATTAATAAAGTGTCAACATCCAGACATTTTCCCATCTTGAGAGATTTCTTAA 733  
 Db 937 taaagaacgaataaaltccaatccaagaataatcttgcacatttggaaagaatcattca 996  
 QY 734 AACCAACCGCTCATCTCTATTTATTCAGATGATGTGATGTAAGCACTTCCACCT 793  
 Db 997 aagcaatgctccactcttgatattatgtcgatgtgagtagtggaagccttccaactct 1056  
 QY 794 TGTCTTGAACAGATTCGTGCTACTTTCAATGCTGTTGCTGCTCAAGCGCCAGATTTG 853  
 Db 1057 tgttctgaacaaagattcgtggaaccttcaacgttagtagcaagtcacacctgttctgg 1116  
 QY 854 TGAATGCTGCAAGCTATGCTTGAAGCAATGCTATCTTGACAGGTGTGACATGTTTAA 913  
 Db 1117 tgaacgtgcgaagaagcattcttgaagaatcgcacattcaagggcggaacagttacac 1176  
 QY 914 AGAGGATCTAGACTGATTAATAAGATGCTCAATGACAGCCCTTGGACAGCTGCTTAA 973  
 Db 1177 agaagaaccttgcttgagtgatgaaagatgcaaatgaaagccttctgtcgaagcagcgag 1236

QY	974	GATTACAGTTGATTAATAATAGCAGACAGTAATTTGTTAGAGTTGACGAAGTTCGAAGCAT	1033
Db	1237	agtagccgtyggacaagaatagaccggtctatctagaaggctgacgaataaccctgaacgcat	1296
QY	1034	TGCTAACCCGATTTTGCACCTGATTTAAATCGCAATTTAGAAACAACACTTTCGACTTTGACCG	1093
Db	1297	ttctccacgcggttgccgttatccaagctccaatccgaacttaacacttccgaatttgaccg	1356
QY	1094	TGAAAACTACACAGACGTTTGGCGAAATTTAGCTGGTGTAGCTGTTATCAAAATAGG	1153
Db	1357	tgaanaattgcaagaacgcttggccaattgtccagtgtgctgagcgttattaagttcgg	1416
QY	1154	AGCTCCACAGACAGACACTTTAAAGAAATGAAGAACTCCGATTGAGAGATGCTCTAAATGC	1213
Db	1417	agccgcacacttgaacacttgatgttgaagaataatgaaacccgcacttgaagaagccctcaacgc	1476
QY	1214	TACACGTCGACGCCCTTGAACAGATATCGTTGCTGGTGGACACGACACTTATATACGT	1273
Db	1477	taccctgycacgctgttgaagaagaatlatctgttcaggctggctggaacacgctcttgccaatgt	1538
QY	1274	TATTGAAAAAGCTACACAGCTTCTTGAGCTTGAAGGCGATGATGCTACGACGTAACATTGT	1333
Db	1537	gattccagctggttgctactaccttggaattgacaggaagatgacgaacgaagcgtatattgt	1596
QY	1334	GCTTCGTGCTCTAGACAGACGCTGTAGCTCAATTTGCTTTAAATGCTGGTACGAAGCTC	1393
Db	1597	tctccgtgcttggagaagaaccgcgtctgcataattgtccacaatgacgaagatttgaagatc	1656
QY	1394	CGTAGTTATTGACACAGTTTGAAAAAACAGCCCTGCAGACAGAGATTAAATGTCGAACAG	1455
Db	1657	tatcgtlatcgaatcgtcttggaaaaatgctgaagcttggctataggatlttaacgcagcaacctcg	1718
QY	1454	TGAGTGGGTTGATTTGATTTAAACAGGAATCATTTGACCCCTGTCAAAGTAACAGATCAGC	1513
Db	1717	cgatgtggttaacactgattgatacgaagatcatctgaccagttaaaatgagctgctcagc	1776
QY	1514	GCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTTTGACACAGAGACAGTGTGCTTAA	1573
Db	1777	ccctacaatatgcagcatctctgtagccagcttgattttgcaacaagaagcagctcgtagccaa	1838
QY	1574	TAAACCTGAAACACAGCTACGCCAGCGCAGCAATGCCAGACAGGTATGAGATCCAGGAATGAT	1633
Db	1837	taaacaccgaacacagtagcccaagctccacgaatgagatccaaacatgatgagcggagatgat	1896
QY	1634	G 1634	
Db	1897	g 1897	
RESULT 4			
X12979			
ID X12979 standard; DNA; 3625 BP.			
X12979;			
XX 19-MAR-1999 (first entry)			
XX Enterococcus faecalis genome contig SEQ ID NO:42.			
XX Enterococcus faecalis; contig; detection; Enterococcal infection;			
RW vaccine; attenuation; computer readable medium; ds.			
XX Enterococcus faecalis.			
OS			
XX W09850555-A2.			
XX			
XX 12-NOV-1998.			
XX			
XX 04-MAY-1998; 98WO-US08985.			
XX			
XX 14-NOV-1997; 97US-0066009.			
XX 06-MAY-1997; 97US-0044031.			
XX 16-MAY-1997; 97US-0046555.			
XX			

XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Barash SC, Dillon PJ, Kunsch CA;  
PI  
XX  
XX MPI: 1999-045171/04.  
DR  
XX  
XX  
XX New isolated Enterococcus faecalis polynucleotides and polypeptides  
PT - used to develop products for the detection of Enterococcus and for  
PT use in vaccines for prevention or attenuation of Enterococcus  
PT infection.  
XX  
XX  
XX Claim 1: Page 414-416; 2084pp: English.  
XX  
XX  
XX A computer readable medium has been developed which has recorded on it  
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.  
CC X12938 to X13919 represent these nucleotide sequences which are primary  
CC nucleotide sequences, also known as contigs. The computer-based system  
CC can identify fragments of the Enterococcus faecalis genome with  
CC commercial importance. The products can be used to detect the presence  
CC of Enterococcus faecalis in samples. They can also be used for  
CC diagnosing Enterococcal infection in an animal and monitoring  
CC progression of disease, and for identifying agents which can be used to  
CC modulate the growth or pathogenicity of Enterococcus faecalis, or  
CC another related organism, in vivo or in vitro. In particular the  
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences  
CC can be used in vaccines to prevent or attenuate an Enterococcal  
CC infection.  
XX  
XX  
XX Sequence 3625 BP; 1335 A; 542 C; 750 G; 993 T; 5 other;

Query Match	56.7%	Score 942.6	DB 20	Length 3625
Best Local Similarity	74.2%	Prod. No. 4.7e-237		
Matches 1191	Conservative	0	Mismatches 415	Indels 0
			Gaps	0
QY	15	ATGCGAAAAAAGAAATCAAAATTTTCACAGATGCGCGCTGCATGGTCGCGAGATTGAT	74	
Db	364	atgycacaaagaaanttaaatcttcgacaabaatgacgtgcgaacatgctcgcgcgagtagat	423	
.QY	75	ATGTTAGCAGATACCGTCAAACTAACGCTTGCTCTAAAGGGCGCAATGTTCTTGAA	134	
Db	424	gatttagcagatacagtagaagtgacattagacctaaagylcgtaacgttgcttagaa	483	
QY	135	AAAGCTTTTGCTTCCTCCCTAATTACTATATGACGGGGTACCATTTGCTAAAGAGATCGAA	194	
Db	484	aaatcatcttggctcacatcttgatactaaagcttgagatacaattgctaagaatctgaa	543	
QY	195	TTAGAGATCATTTTGGAAAAACATGGGAGCAAAATTTGGTGTGAAGTGCGCTTCAAAAC	254	
Db	544	ttggaaagatcatcttggaaaaacatgycgcacaaatlagttcgaagcttgctctcctaaagc	603	
QY	255	AATGATATTGCTGCTGATGGAGCGAGCTACTGCAACAGTTTTCACACAAGCCATTGGTCAT	314	
Db	604	aatgatattgctgcygacggaacacacacagagactgtcttgacacaaagccattgctg	663	
QY	315	GAAGGACGTAAAAAATGTGACAGCAGCTGCTAATCCAATTTGGTATCCGTGAGGCAATTGAA	374	
Db	664	gaaggtctttaaataaagcttaactgcttgagacacacacattagatcttcgctgctgagtag	723	
QY	375	ACAGACACGCAACGACGCTTTGACGCTTTGAAAGCATTGGTCTCAACGCTGTATCTGGCAG	434	
Db	724	ttagcaacacaaacacgcgctagaaagaattacacaaatttcaatctgttgattccaataa	783	
QY	435	GAAGCTATTGCTCAGGTCGCTGCTGACATATACACGCTGTAAAAAGTTGAGAGATATAT	494	
Db	784	gaagcatctgcaacagtcgctgctgttcaatcaggtcttgaaaagctcgcccaattaat	843	
QY	495	TCAGAAAGCTATGAGACGCTGTGGGCAACGATGGTGTGATTACATGGAAGATCTCGAGCT	554	
Db	844	gccgactgcaatgtgtaaaaatgttgtaacagcgcgttaattacattgaaatctcaaaaag	903	
QY	555	ATGGAACCGAAGCTTGAAGTGGTTCAAGGCAATGCAATTTGACCTGGTGTACTGCTCTGA	614	

## DESCRIPTION

Query Match	Similarity	42.4%	Score 704.2	DB 18	Length 1647
Best Local	Similarity	66.4%	Pred. No. 8.2e-175		
Matches 1044	Conservative	0	Mismatches 523	Indels 6	Gaps 2
QY	19	CAAAAGAAATCAAAATTTTTCAGCAGATGCGCGTGCATGGTGGCGGAGTTGATATGT	78		
Db	8	ctaaagaatcctctttgatgtctaaagccgcygaaaactcttcacgaggtgtgaataac	67		
QY	79	TAGCAGATACCGTCAAAATACGCTTGGTCCCTAAAGGGCGCAATGTGTTCTTGAAG	138		
Db	68	ctgcaaaagcgtctaaagtaacaacttggaaaccaaagccgtaattctgctatctgaaagt	127		
QY	139	CTTTTGATTCCCTTAATTAATAATGACGGGGTAACCATTTGCTTAAAGATCGAATTAG	198		
Db	128	ctttgttcccccagtaattacaagaagtgtgtatctctgtgcaagaagaatttgacttg	187		
QY	199	AAAGTCAATTTGAAAAACATGGGAGCAAAATTTGGTCTCGAAGTGGCTTTTAAACCAATG	258		
Db	188	aagtaagatttggaaaatttggcgtctcaaatgggtctaaagtaagctctcccaactatgcg	247		



OY	259	ATTTCGGGATGGAGGACACATACGCAACAACTTTTGGACAAAGCACTGTTCAATG	318
Db	248	atatctgcgtgtagtgaacacacacacagcagctctgcacaagctatattacgttag	307
OY	319	GACTAAATAATGTGCACGAGCGTCTAATCCAAATTTGGTATCCGTCGAGGCAATTGAACG	378
Db	308	gtgtaaacctgttagaagcgtgcgttaacccatagccatataaogtgcataagtaag	367
OY	379	CAACAGCAACAGCTGTTGGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAGAG	438
Db	368	ctgtctgtcgtcttactaaagaactaagcagcatlacaagaactactcgtgcacaaag	427
OY	439	CTATTGCGAGGTGGCGGCASTATCATCAGCGCTGA---AAAAGTTGAGAGGTATATCT	495
Db	428	aaatagcctaaagttggaaacatcttcgtcaaacctctgtacacaatagtaatacatag	487
OY	496	CAGAGAGCTATGAGACCGTGTGGGCAACAGATGNTGTCAATACCATCAAGAAATTCAGGTA	555
Db	488	ctgaagctatggtcctaaagttggaaagaggtgtatcacagttgaggaagctaaagtc	547
OY	556	TGGAACAGAACTTTGAAGTGGTTGAAGGCATGTCATTTGACCGTGGTTAACCTGTCAT	615
Db	548	ttgaactaactatgtagtgcgttgaaagatgagttgacggtgcgtactcctccat	607
OY	616	ACATGTGTACAGACATGAAAAAATGTTGCAGACCTTGAAAAACCATTTATCTTAATCA	675
Db	608	acttctgaactaaacctgtgagaaatagtgtgtgaacttataccttataccttactcgtg	667
OY	676	CGATTAATAAAGGTTCCAAACATCCAGACACTTTTGGCAGTACCTTGAGGAAGTTCTTAA	735
Db	668	atgtagaaaagatctactaagcatgaagacatgctaccaactcttagaacaagttgcataag	727
OY	736	CCAAAGCTGCATTACTCATTTATTTGCAGATGATGTGATGATGTAAGACACTTCCAACTTTG	795
Db	728	taaacgcgtccactctattattctgcgaagagctagaaggtggaagcaactgtaaacactgt	787
OY	796	TCTTGAACAGAGTTCGTGACTTTCAATGTGTTGCTCTCAAGGCGCAGGATTTGGTG	855
Db	788	tagtacaataagctccgcgtgagacactcaagttgtagcgtlaaaagctccgtgtttgtgt	847
OY	856	ATTCCTCGTAAGAAGTATCTTGAAGACATGTCATCTCTGACAGTGGTACAGATTTACAG	915
Db	848	aagcgcgttaaagctatgcgttgaagaatctgcatccctacgcgtgagagagaacataattg	907
OY	916	AGGATCTAGAGACTTGAATTTAAAGATGCTAACATGACAGCCCTTGAGACAGCTGTACAGA	975
Db	908	aagtcgtgtagtaaaagctgtgaataatgtaaagctgtcctctcttaggaacagtaaacgtg	967
OY	976	TTACAGTTGTAATAAGATAGACACAGTATTTGTTAGAGTTACAGAAAGTTACAGAGCTATTG	1035
Db	968	tagtattatgtaaaagaatactactactcgtgtatgtgtcgtgaataatagaagatactat	1027
OY	1036	CTAACCGGATTTGGCATTCGATTAATTCGAATTTGAACCAACAACCTCTGACTTTGACCGTG	1095
Db	1028	aagctcgtagttaaacaaatctgcgcacaacttgaagaacaaagctcagatattgctgtg	1087
OY	1096	AAAACTACACGAACGTTTGGCCAAATTTAGCTGTTGTTAGCTGTTATCAAGTAGAGAG	1155
Db	1088	aaaaactcaagaacgctctgcgaacaactcgttggtagagtagctgtatcacatgtagag	1147
OY	1156	CTCCACACAGACAGCTTTAAAGAAATGAACCTTCGATTTGAGAGATGCTTAATATGTA	1215
Db	1148	ctgtactactgaactcgtaaatgaagaagatcgtgttagaagaatgctcttaaatgtcaa	1207
OY	1216	CAGCTGACGCGCTTGAGAGAGGTATGTTGTCGTGTGTGTGTGAGACACACTTATTTACGGTTA	1275
Db	1208	caagagcgtgcgttggaaagagtagtgcctcgtgtagtgtagtactcttctgtgcgcctca	1267
OY	1276	TTTGAATAAAGTAGAGCTTTGAGCTTGAGAGGCACTATG---CTAATGGACCTACATTTG	1332
Db	1268	tttaagtcctctgtatgatatctaaacctcgtatgatagtagaigtaactcgtgcgtactaataca	1327
OY	1333	TGCTTGTGCTCTTGAAGAGCCTGTACGTCAATTTGCTTTAAATGCTGGGTACGAAGCT	1392

[illegible]





OS	Streptococcus pyogenes.
XX	
PN	WO9935270-A1.
XX	
PD	15-JUL-1999.
XX	
PF	29-DEC-1998; 98MO-CA01203.
XX	
PR	31-DEC-1997; 97US-0001737.
XX	
PA	(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX	
PI	Mizzen L, Wisniewski J;
XX	
DR	WPI; 1999-430397/36.
DR	P-PSDB; Y23903.
XX	
PT	New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT	useful in vaccines, as carriers for other immunogens, as anticancer
XX	agents and for diagnosis
XX	
PS	Claim 3; Fig 3A-B; 176pp; English.
XX	
CC	The present sequence encodes a heat shock protein, designated Hsp60-1.
CC	The protein, its fragments, variants and fusion proteins, are
CC	used to elicit or enhance an immune response against Streptococcus,
CC	and to elicit a similar response to a target antigen fused to the
CC	protein. Unlike other immunological carriers, Hsp60 proteins are not
CC	immunosuppressive so provide an increased response to any conjugated
CC	fused antigen. Also, where used for cancer control, they lack the side
CC	effects associated with endotoxins. They can also be used to detect
CC	specific antibodies and in treatment or prevention of tumours
CC	(e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC	liver). The Hsp60 polynucleotide is used for recombinant production
CC	of the protein, as a source of primers and probes for detecting
CC	streptococci in standard hybridization/amplification assays, and
CC	therapeutically in gene therapy vectors.
XX	
Sequence	1662 BP; 429 A; 414 C; 498 G; 321 T; 0 other;

Dd	1501	caaccacaaagcggccgcttcgtcctcgagtaacgcgcacatctgtcgtcgtgatcatca	1560
Oy	1552	CACACAGACGACGATTGTTCCTAATAAACCGTCAACCCAGCTACGCCGAC	1596
Dd	1561	ccaccgagtgatcgttgcaccgacctgcctaagaagcacgacctgt	1605
RESULT	9		
ID	Z54509	standard; DNA; 1635 BP.	
XX			
AC	Z54509;		
XX			
DT	21-MAR-2000	(first entry)	
DE	Neisseria meningitidis ORF 982 partial DNA sequence SEQ ID NO:2965.		
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;		
KM	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;		
XX	antibacterial; gene therapy; ds.		
OS	Neisseria meningitidis.		
XX			
PN	MO957280-AZ.		
XX			
PD	11-NOV-1999.		
XX			
PF	30-APR-1999;	99WO-US09346.	
XX			
PR	01-MAY-1998:	98US-0083758.	
PR	31-JUL-1998:	98US-0094869.	
PR	02-SEP-1998:	98US-0098894.	
PR	02-SEP-1998:	98US-0099062.	
PR	09-OCT-1998:	98US-0103749.	
PR	09-OCT-1998:	98US-0103794.	
PR	09-OCT-1998:	98US-0103796.	
PR	25-FEB-1999:	99US-0121528.	
XX			
PA	(CHIR ) CHIRON CORP.		
PA	(GENO-) INST GENOMIC RES.		
XX			
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;		
PI	Petersen J, Pizsa M, Rappoli R, Ratel G, Scalato E, Scarselli M;		
PI	Tetelin H, Venter JC;		
XX			
DR	WPI; 2000-062150/05.		
DR	P-PDB; Y75747.		
PT	Novel Neisserial polypeptides predicted to be useful antigens for		
PT	vaccines and diagnostics		
PS	Claim 7; Page 1387-1388; 1453pp; English.		
XX			
CC	Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent		
CC	novel Neisseria meningitidis and N. gonorrhoeae polynucleotides and		
CC	polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR		
CC	primers used in the exemplification of the present invention. The		
CC	polypeptides, the polynucleotides, antibodies and compositions of		
CC	the invention can be used as vaccines, as diagnostic reagents, and as		
CC	immunogenic compositions. The polypeptides can be used in the		
CC	manufacture of medicaments for treating or preventing infection due to		
CC	Neisserial bacteria (e.g. meningitis and septicaemia), to detect the		
CC	presence of Neisseria bacteria, or to raise antibodies. They may also		
CC	have use as screen for agonists or antagonists, which may themselves		
CC	may also be used in gene therapy protocols.		
XX			
SQ	Sequence 1635 BP; 448 A; 440 C; 456 G; 291 T; 0 other;		
Query Match	33.3%; Score 552.8; DB 21; Length 1635;		
Best Local Similarity	60.0%; Pred. No. 3.4e-135;		

[illegible]



```

Db 221516 TTTTAGGCAAAATAACTTTATTAATTTATCTGAGAGATATTAGGGGAGCTCTTG 221457
QY 787 CAACCTTGTCTTGAACAGATTCGTGTACTTTCAATGTGGTGTCTGTCAAGCGCAG 846
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221456 CTGCTCTTGTATTAACAGGTTAGAGAGCTTTAAAGTATGCAATTAATCTCTG 221397
QY 847 GATTTCGTATCGTCGTAAGCTATGTCTTGAAGCATTTGCTATCTTACAGCTGTACAG 906
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221396 GTTTTGTGTAGACGAAAGCAATGCTTGAGATATTGCAAGTCTTACCGGCGTGT 221337
QY 907 TGATTACAGGATCTAGACTTAATTAAGATGCTACATACAGCCCTTGACAGG 966
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221336 TAATCATGTGAGAGCTAGGCTTACTCTGTGACAGACTGAATGAACCACTTGACAGG 221277
QY 967 CTGTAAGATTACAGTTGATTAAGATAGACAGATTAATTTGTAAGTTTCAAGAGTTACAG 1026
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221276 CTAAACTATTAAAGTTGATTAACATACACTATTATTAA---TACCGGCAATAAG 221220
QY 1027 AAGCTATTGCTAACCGTATTGCACTGATTAATTCGCAATTGAAACAACAACCTTGACT 1086
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221219 AGCAATTAAGAGAGCTTCAAGAGCTTATTAAGCAATTAAGATTAACATCTGAT 221160
QY 1087 TTGACCGTGAATAAATCTAAGAAAGTTGGGAAATTAAGCTGGTGTGTACTGTATCA 1146
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221159 ATGATTAAGAAAACTTCAAGAGCGCTTGCATAAAGCTTGGGAGGTGCTGTATTA 221100
QY 1147 AAGTAGAGCTCCAAACAGACAGACCTTTAAAGAAATGAATTCGCAATTGAGGATGCTC 1206
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221099 ATGTGAGAGCTGTACTAGAGTAGAGCTTAAGGAAAAAGCAATAGAGTTAGAGACGCTC 221040
QY 1207 TAAATGCTACAGCTGCAAGCGCTGTAAGAGATCGTTGCTGTGTGTGGAACAGCACTTA 1266
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221039 TTCTGCAACTCGTGTGTGTGTGTAAGAGGCTTGTGCTGGCGGTGATCACTCTTA 220980
QY 1267 T-----TACGTTATTGAAAAAGTAGACGCTTGTGAGCTTGAGGCGCATATGCTA 1317
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220979 TTGAAGTTGCTATGTATTTAGATCAATAGATACAGTAATTAAGCTATGAGAAAAAGC 220920
QY 1318 CTGACGATACATATGCTGCTGCTAGAGAGCCGTGACGTAATTCCTTTAATG 1377
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220919 AAGGTTTGTAGATGTAAAAAGAGTCTTGAGAGCCCATAGACAGATTTTCCAATG 220860
QY 1378 CTGGGTACGAAGCGCTCCGTAGTTTGAACAAGTGAAGAAAAAGCCCTGCAGAAACAGAT 1437
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220859 CTGGTTTGAAGGATCTATTTATTCATCAATTAAGCAAAAAAGAGGCTTGGGT 220800
QY 1438 TTATTCGTGCAACAGAGTGGGTGATATGATTAAGCAAGATATGACCTGTCA 1497
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220799 TTGATGCTTCCAGCTTTTAAGTGGGTAATATGATTTGAGAGTAATATGATCTCTGTA 220740
QY 1498 AAGTAACAGATCAGCGCTCAAAATGACGCTTGTAGCTAGTCTTATTTTGAACACAG 1557
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220739 AAGTTACAAAGAGCGCTTCAAAATCTGCTTAATTTGCTGGACTTTTATTAACACAG 220680
QY 1558 AAGCAGTTGTTGTAATTAACCTGAACAGCTACGCCAGCGCCAGCAATGCCACAGGTA 1617
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220679 AATGTGAATCAGATATTAAGAAAGAAAAATACTTTCGTGTGTGTGTGTATCTCTA 220620
QY 1618 TGAATCCAGAAATG 1631
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220619 TGGACCCAGGAATG 220606
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

## RESULT 11

Q13136 ID Q13136 standard; DNA; 2465 BP.

XX AC Q13136;

XX DF 22-OCT-1991 (first entry)

XX DE Hyp operon.

XX

```

KW Antibodies; heat shock; hypersensitive; allergen; HSP60; GroEL;
KW GroES; ss.
XX
XX Chlamydia psittaci GPIC.
XX
FH key Location/Qualifiers
FT CDS 314..622
FT     /tag= a
FT     /label= hypA
FT     673..2307
FT     /tag= b
FT     /label= hypB
FT     83..88
FT     -35_signal
FT     /tag= c
FT     120..127
FT     /tag= d
FT     301..304
FT     /tag= e
FT     656..659
FT     /tag= f
FT     2404..2440
FT     terminator
FT     /tag= g
FT     /note= "dyad symmetry; rho-independent"
FT     350..385
FT     /tag= h
FT     /label= oligo
FT     /note= "used for Northern blot analysis"
FT     2152..2187
FT     /tag= i
FT     /label= oligo
FT     /note= "used in Northern blot analysis"
PN U5751317-A.
XX
XX 09-JUL-1991.
PD
XX
XX 31-MAY-1990; 90US-0143560.
PR
XX
XX 31-MAY-1990; 90US-0531317.
PR
XX
XX (USSH ) NAT INST OF HEALTH.
PA
XX
XX WPI: 1991-245693/33.
DR
XX
XX P-PSDB: R13334.
DR
XX
XX P-PSDB: R13335.
XX
PT DNA encoding HypA and HypB Chlamydia proteins - used to develop
PT prods. for detection of and vaccines against Chlamydia infection.
XX
XX
XX Disclosure; Fig 5; 51pp; English.
PS
XX
XX The sequence was obtd. from clone pop57 isolated from a DNA library
XX prepd. from C. psittaci genomic DNA. The two ORFs encode HypA and
XX HypB hypersensitivity proteins, regulated by a heat shock type
XX promoter region. The DNA can be used to produce recombinant
XX proteins or to design probes for the detection of chlamydial
XX infection. The Hyp proteins can be used to to raise antibodies and
XX in vaccines. The Hyp B protein elicits a cell-mediated immune
XX response so can be used as a skin test antigen.
XX See also Q13137.
XX
XX Sequence 2465 BP; 853 A; 467 C; 509 G; 636 T; 0 other;
SO

```

## Query Match

Best Local Similarity 33.1%; Score 550; DB 12; Length 2465;

Matches 951; Conservative 0; Mismatches 615; Indels 12; Gaps 2;

```

QY 18 GCAAAAGAAATCAATTTTACGAGATGCGCTGTCGCAATGAGTGGCGGAGTGTATG 77
DB 679 gcaaaaataatcaataacgaagacgcagaaaaaaacccataaaggaggttaaaccc 738
QY 78 TTACGATACCGTCAAAAGTAACGCTTGCTCTAAGGCGCAATGTTGTTGAAAAA 137

```

Db 739 cttgcgaagaagctgtaaaagacccttagtccctaaagaagccgctcaatgctgctacgtacaa 798  
 QY 138 GCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCAATGGCTAAAGAGATGAAATTA 197  
 Db 799 agcttctgctcctcaagctacaaagatgctgctacgtctgctacgaagaatctgctc 858  
 QY 198 GAAGATCATTTTAAACATGGAGCAAAATTTGGTGTCTGTAAGTGGCTTTAAACCAAT 257  
 Db 859 gaagacaagatgtagaagctggagctcaatctgtaaaagaagctgctgacaaactgca 918  
 QY 258 GATATTGCTGTGATGGAGCACTACTGCAACAGTTTGGACACAGCAATGTTGCAATGAA 317  
 Db 919 gataaagctggtagtgaactacacaagctacgtctctctgcaagaagctacacagtgaa 978  
 QY 318 GGATTAATAAATGTGACAGCAGCTGCTAATTCATTTGTATCCGTGAGCAATGGAACA 377  
 Db 979 gggcttgagaagcttaactgtagagcgccatccctacgtacctaagaagcaltgataag 1038  
 QY 378 GCACAGCAACAGCTGTGAAAGCTTGAAGCCATTGCTCAGCACTGTATCTGGCAAGGAA 437  
 Db 1039 gcaagtaaaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1098  
 QY 438 GCTATTGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494  
 Db 1099 gaataagctcaagctgaagctacatctctgcaataatgctgctgctgctgctgctgctgct 1158  
 QY 495 TCAGAGCTATGAGAGCGTGTGGCAACGATGCTGTGATTAACATTCAGCAAGATCTCGAGT 554  
 Db 1159 ggcgaagcaltgtaaaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1218  
 QY 555 ATGGAACAGCACTTGAAGTGTGAAAGCATTTACCGTGTGTTACCTGCTCA 614  
 Db 1219 ttcgaagctcctcgcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1278  
 QY 615 TACATGCTCAGACATGAAATGTTGACAGCTTGAACCAATTTATCTTAATC 674  
 Db 1279 tactcttcaacaatcctgtaaaacaagaatgctgctttagaagaagctcgtgcttacc 1338  
 QY 675 ACGGATTAATAAAGCTGTCAAAACATTCATTTGCCACTACTTGAAGAACTTTTAA 734  
 Db 1339 tatgataaaaaatctcgcgaatcaagaatcttctacccgcttctacaaagaagtagaa 1398  
 QY 735 ACCAACCCTTCATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794  
 Db 1399 tcaagagctcctcctacttactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1458  
 QY 795 GCTTGAACAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 854  
 Db 1459 gtagtcaaaagctacgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1518  
 QY 855 GATCGTGTAAAGCTATGCTTGAAGCATTTGCTATCTTGAACAGTGTACAGTATTA 914  
 Db 1519 gatatacaaaaagctatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1578  
 QY 915 GAGAGCTGAGCTTAATTAATAAGTGTACATGACACCTTTGACAGCTGCTTAAG 974  
 Db 1579 gaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1638  
 QY 975 ATTACAGTGAATTAAGATGACAGATTAATTTGAGAGTTCAGAGTTCAGAGTATTA 1034  
 Db 1639 gtcactgcttccaaagaataatacaaatctgctgctgctgctgctgctgctgctgctgctgctgct 1698  
 QY 1035 GCTAACCGTATTTGCATGATTAATTCGAATTAAGAAACAACAACCTTCTGACTTTGACCGT 1094  
 Db 1699 gaatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1758  
 QY 1095 GAAAAACTCAAGACGTTGGCGAAATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1154  
 Db 1759 gaataactcaagaagctttagctaaacttccgagcgctgctgctgctgctgctgctgctgctgct 1818  
 QY 1155 GCTCCACAGAGAGCTTTAAAGAAATGAAACTGCTGATGAGGATGCTGCTAAATGCT 1214

Db 1819 gctgctacagaatcgaaatgaaagaagaaaaagacagagtagatgctgctgctgca 1878  
 QY 1215 ACAGTCCACCCGTTGAAGAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274  
 Db 1879 acctctgct 1938  
 QY 1275 ATTGAATAATGACAGCTCTTGAAGCTTG-----AGGCGATGATGCTACTGACGT 1325  
 Db 1939 atccctactttagaagcttccatctcctacttctacaaatgaagatgagcaaatcgagca 1998  
 QY 1326 AACATGCT 1385  
 Db 1999 cgtatgcttccaaagataltccgcctcctacatlaagaacatgcaagcaatgctgctgctgct 2058  
 QY 1386 GAAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445  
 Db 2059 gaagcgctacatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2118  
 QY 1446 GCACAGGCTGAGTGGTGTGATGATTAATAACAGAAATCAATTAACCTGTCAAGTAA 1505  
 Db 2119 ttagcgatgcttcaacacgcatgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2178  
 QY 1506 CGATCAAGCGCTTCAAAATGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565  
 Db 2179 cgtctgctttagaagaagcagcagcttctgctgctgctgctgctgctgctgctgctgctgctgct 2238  
 QY 1566 GTTGCTAATTAACCTGAA 1583  
 Db 2239 atgctgctatctctgaa 2256  
 RESULT 12  
 A13000  
 ID A13000 standard; DNA; 1635 BP.  
 AC A13000;  
 DT 15-AUG-2000 (first entry)  
 DE  
 XX Neisseria meningitidis gene sequence SEQ ID NO:2964.  
 DE  
 XX  
 DE  
 KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy; ds.  
 OS  
 XX Neisseria meningitidis.  
 OS  
 XX  
 PN MO9957280-A2.  
 PD 11-NOV-1999.  
 XX  
 XX 30-APR-1999; 99WO-US09346.  
 PF  
 XX  
 XX  
 PR 01-MAY-1998; 98US-0083758.  
 PR 31-JUL-1998; 98US-0094869.  
 PR 02-SEP-1998; 98US-0096994.  
 PR 02-SEP-1998; 98US-0099062.  
 PR 09-OCT-1998; 98US-0103749.  
 PR 09-OCT-1998; 98US-0103794.  
 PR 09-OCT-1998; 98US-0103796.  
 PR 25-FEB-1999; 99US-0121528.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 PA  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
 PI Petersen J, Pizsa M, Rappelli R, Ratti G, Scatato E, Scarselli M;  
 PI Tettelin H, Venter JC;  
 PI  
 DR WPI: 2000-062150/05.  
 XX  
 XX  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics













**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:54:05 ; Search time 71.35 Seconds  
(without alignments)  
3751.749 Million cell updates/sec

Title: US-09-001-737-7  
Perfect score: 1661  
Sequence: 1 GAATCGCTTCATATGCA.....TGGCGGATAGCCGAATTC 1661

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/PTCUT\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.6	31.8	1838	3 US-08-470-260-7	Sequence 7, Appli
2	528.6	31.8	1838	3 US-08-471-491-7	Sequence 7, Appli
3	528.6	31.8	1838	3 US-08-466-662-7	Sequence 7, Appli
4	496.6	29.9	2284	2 US-08-467-822-28	Sequence 28, Appl
5	446.8	26.9	1626	2 US-08-997-080-159	Sequence 159, App
6	446.8	26.9	1626	2 US-08-997-362-159	Sequence 159, App
7	446.8	26.9	1626	3 US-09-095-855-159	Sequence 159, App
8	444.8	26.8	1569	2 US-08-997-080-113	Sequence 113, App
9	444.8	26.8	1569	2 US-08-997-362-113	Sequence 113, App
10	444.8	26.8	1569	3 US-09-095-855-113	Sequence 113, App
11	416.6	25.1	1620	2 US-08-461-775-10	Sequence 10, Appl
12	416.6	25.1	1620	3 US-09-031-606-10	Sequence 10, Appl
13	411	24.7	2668	2 US-08-461-775-11	Sequence 11, Appl
14	411	24.7	2668	3 US-09-031-606-11	Sequence 11, Appl
15	362	21.8	1320	2 US-08-461-775-8	Sequence 8, Appli
16	362	21.8	1320	3 US-09-031-606-8	Sequence 8, Appli
17	353.2	21.3	2167	2 US-08-461-775-9	Sequence 9, Appli
18	353.2	21.3	2167	3 US-09-031-606-9	Sequence 9, Appli
19	273.8	16.5	985	2 US-08-997-080-161	Sequence 161, App
20	273.8	16.5	985	2 US-08-997-362-161	Sequence 161, App
21	273.8	16.5	985	3 US-09-095-855-161	Sequence 161, App
22	271.8	16.4	927	2 US-08-997-080-116	Sequence 116, App
23	271.8	16.4	927	2 US-08-997-362-116	Sequence 116, App
24	271.8	16.4	927	3 US-09-095-855-116	Sequence 116, App
25	178.4	10.7	568	3 US-08-714-918-51	Sequence 51, Appl
26	176.6	10.6	647	2 US-08-997-080-115	Sequence 115, App
27	176.6	10.6	647	2 US-08-997-362-115	Sequence 115, App
28	176.6	10.6	647	3 US-09-095-855-115	Sequence 115, App

Query Match 31.8%; Score 528.6; DB 3; Length 1838;  
Best Local Similarity 58.9%; Pred. No. 1.4e-146;

29 101.2 6.1 343 2 US-08-473-020A-7 Sequence 7, Appli  
30 100.4 6.0 337 2 US-08-997-080-77 Sequence 77, Appl  
31 100.4 6.0 337 2 US-08-997-362-77 Sequence 77, Appl  
32 100.4 6.0 337 3 US-08-873-970-77 Sequence 77, Appl  
33 100.4 6.0 337 3 US-09-095-855-77 Sequence 77, Appl  
34 97.4 5.9 350 1 US-08-105-168B-22 Sequence 22, Appl  
35 97.4 5.9 350 2 US-08-698-948-22 Sequence 22, Appl  
36 94.8 5.7 342 2 US-08-473-020A-31 Sequence 31, Appl  
37 94.8 5.7 343 2 US-08-473-020A-4 Sequence 31, Appl  
38 91.8 5.5 343 2 US-08-473-020A-2 Sequence 2, Appli  
39 91.6 5.5 342 2 US-08-473-020A-28 Sequence 28, Appl  
40 91.6 5.5 342 2 US-08-473-020A-30 Sequence 30, Appl  
41 91.6 5.5 343 2 US-08-473-020A-3 Sequence 3, Appli  
42 90.2 5.4 342 2 US-08-473-020A-29 Sequence 29, Appl  
43 90 5.4 360 2 US-08-997-080-79 Sequence 79, Appl  
44 90 5.4 360 2 US-08-997-362-79 Sequence 79, Appl  
45 90 5.4 360 3 US-08-873-970-79 Sequence 79, Appl

## ALIGNMENTS

RESULT 1  
US-08-470-260-7  
; Sequence 7, Application US/08470260  
; Patent No. 6077706  
; GENERAL INFORMATION:  
; APPLICANT: Covacci, Antonello  
; APPLICANT: Bugnoli, Massimo  
; APPLICANT: Telford, John  
; APPLICANT: Macchia, Giovanni  
; APPLICANT: Rappuoli, Rino  
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful  
; TITLE OF INVENTION: for Vaccines and Diagnostics  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/470,260  
; APPLICATION NUMBER: US/08/470,260  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,848  
; FILING DATE: 21-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0316.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1838 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-470-260-7

Matches 966; Conservative 0; Mismatches 664; Indels 9; Gaps 3;

Qy	13	ATATGGCAAAAGAAATCAAAATTTTCACAGATGCGGGTGCTGCCATGCTGCGCGGAGTTG	72
Db	56		
Qy	73	ATATGTTTGGCAGATACCGTCAAAAGTAACCTTTGGTCCCTAAAGGGCGCAATGTTGTTCTTG	132
Db	116	GGCAACTCCATGACGTGTCTAAAGTAACCATGGGCGCCAAAGAGGCAAGGATGATTGATGCC	175
Qy	133	AAAAAGCTTTTGGTGTCTCCCTTAATTTACTTAATGACGGGTAAACCATGCTCAAGAGATCG	192
Db	176	AAAAAGCTATGGCGTCCCAAGCATCACCAAGACGGCTGAGCGTGCTTAAAGAGATTG	235
Qy	193	AATTAGAAGATCATTTTGAACAACATGGGAGCAAAATTTGGTGTCTGAAGTGGCTTCAAAA	252
Db	236	AATTAAGTTGCCCAAGTAGATTACATGGCGGCTCAACTCGTTAAAGAAAGTAGCGAGCAAAA	295
Qy	253	CCAATGATATTGCTGTGTGATGGGACGACTACTTGCACACAGTTTGTGACACACCCATGTTCT	312
Db	296	CCGCTGATGCTCGCGGATAGGCACGACCACACGACCGCTGCTAGCTTATAGCATTTTATA	355
Qy	313	ATGAAGGACTTAAAAATGTGACAGCAGGTGCTTAATCCAAATTTGGTATCCGTCGAGCGCATG	372
Db	356	AAGAAGCTTTGAGGAATATCACGGCTGGGGCTTAACCCATTGAAGTGAACACGAGCATGG	415
Qy	373	AAACAGCACAGACACAGCTGTTTGAAGCCTTGAAGACCATTTGCTCAACGCTGTATCTGGCA	432
Db	416	ATAAAGCTGCTGAAGCGGATCATTAATGAGCTTTAAAAAGCGAGCAAAAAAAGTAGCGGTA	475
Qy	433	AGGAAGCTATTGCTCAGGTGCGTGCAGTATCATCACGGTCT--TGAAAAAGTTGGAGAGT	489
Db	476	AAGAGNAATCACCAAGTGGGACCATTTCTGCAAACTCCGATCNCANATTCGGGAAC	535
Qy	490	ATATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGGTGTATTAACCATCGAAGAATCTC	549
Db	536	TCATCCGTGACGCTATGGAAAAAAGTTGGGTAAAGACGGGTGATCACCGCTTCAGGAAGCTA	595
Qy	550	GAGGTATGGAACAGAACTTTGAAGTGGTTGAAGGATGCAATTTGACCGTGGTTACCTGT	609
Db	596	AGGGCATTTGAAGATGAATTTGGATGTGCTGAAGAAGGATGCAATTTGATAGAGGCTACCTCT	655
Qy	610	CTCAATACATGGTTCACAGACAATGAAAAATGGTTGCAGACCTTGAAAACCCATTTATCT	669
Db	656	CCCCATTATTTGTAAAGCAAGCTGAGAAANATGACCGCTCAATTTGGATTAATGCTTACATCC	715
Qy	670	TAATACGGATAAAAAAGTGTCAAAATCCAAAGACATTTTGCCACTACTTGAAGGAAGTTC	729
Db	716	TTTTTAACGGATAAAAAATCTCTAGCATGAAAGACATTTCTCCGCTACTAGAAAAAACCA	775
Qy	730	TTAAACCAACCGTCCATTACTCATTTATGACAGATGATGTGGATGGTGAAGCACTTCCAA	789
Db	776	TGAAGAAGGGCAAAACCGCTTTTAATCATCGCTGSAAGACATTTGAGGGCGAAGCTTTTAA	835
Qy	790	CCCTGTGCTTGAACAAGATTCTGTGGTACTTTTCAATGTGGTTGCTGTCAAAAGCCGAGAT	849
Db	836	CTCTAGTGGTGAATAAATTAAGAGCGGTGTTGAATATCGACGGTTAAAGCTCCAGGCT	895
Qy	850	TTGGTGATCGTGAAGCATGTCTGTTGAAGACATTTGCTATCTTGACAGTGGTACAGTGA	909
Db	896	TTGGGACAGAAAGAAAAATGCTCAAAAGACATCGCTATTATTTAACCGCGGTCAAGTCA	955
Qy	910	TTACAGGGATCTAGGACTTGAATTAAGAATGCTTACAATGACAGGCCCTTGGACAGGCTG	969
Db	956	TTAGCGAAGAAATTTGGCGTTTGAGTCTAGAAAACGCTTGAAGTGGAGTTTTTATAGGCA	1015
Qy	970	CTAAGATTACAGTTGATAAAGATAGCAGATTAATTTGTTGAAGTTTCAGAAAGTTTCAGA	1029
Db	1016	GAGGATTTGATTGCAAGAAGCAACACCCAGTCGTAGATGGCAAGGCCATAGCCATG	1075
Qy	1030	CTATTGCTAACCGTATTTGCACTGATTAATCGCAATTTAGAACAACAACTTCTGACTTTG	1089
Db	1076	ATGCTTAAAGACAGAGTTCGGCGAGATCAAAACCCAAATTTGCAAGTACGACAGCGATTA	1135

## RESULT 2

```

US-08-471-491-7
; Sequence 7, Application US/08471491B
; Patent No. 6090611
; GENERAL INFORMATION:
; APPLICANT: Antonello
; APPLICANT: Covacci, Massimo
; APPLICANT: Bugnoli, John
; APPLICANT: Telford, Giovanni
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
; TITLE OF INVENTION: Diagnostics
; FILE REFERENCE: CHIR0044
; CURRENT APPLICATION NUMBER: US/08/471,491B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1838
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-08-471-491-7

```

Query Match	31.8%	Score 528.6;	DB 3;	Length 1838;
Best Local Similarity	58.9%	Pred. NO. 1.4e-146;		
Matches 966;	Conservative	0;	Mismatches 664;	Indels 9;
				Gaps 3;

**Qy** 13 ATATGGCAAGAATAATTTTCAGCAGATCGCGGCTGCCATGGTGC GCGGGAGTTG 72  
+ ||||| | ||||| | ||||| + | | | |  
**Db** 56 aaatgcaaaagaatacaaattttcagatatcgcgaacacctttatttcaaggctga 115  
+ ||||| | ||||| | ||||| + | | | |



QY 73 ATATGTTACGATACCGTCAAGTAACGCTTGGTCTTAAGGGCGCAATGTTGTTCTT 132  
 Db 116 ggaactccatgacgctgtcaagtaaccatgggccaagagcaggaatgtattgatcc 175  
 QY 133 AAAAGCTTTTGGTCTCCCTTAATTAATCAATGACGGGTAAACCATTTGCTAAAGAGATCG 192  
 Db 176 aaaaagctatggcgtcccaagcatcaccaaagacgagcgctgagcgctgcaagagatg 235  
 QY 193 AATTAGAAGATCATTTTGAAGACATGGGACAAATTTGGTGTCTGAAGTGGCTTCTAAAA 252  
 Db 236 aattaagtggccagtagctaaactaagcgctcaactcgltaaagaagtgcagagcaaaa 295  
 QY 253 CCAATGATATTTGGTGTGATGGGACGACTACTGCACAGTTTGTGACACAGCCATTTGTTTC 312  
 Db 296 ccgctgatgctgcgagtgagcagcaccacagcgcggtgctagcttatagcattttta 355  
 QY 313 ATCAAGGACTAAAAATGTGACAGCAGGTGCTAATCCAATTTGTATCTCCGTCGAGGCAATTG 372  
 Db 356 aagaaggttgaggaataatacagcgctggggctaacctcttgaagtgaacgagggcatgg 415  
 QY 373 AAACAGCAACAGCAACAGCTGTTTGAAGCCCTTGAAGCCATTTGCTCAACCTGTATCTGGCA 432  
 Db 416 ataaagctgctgaagcagatcatttaagtgcgttaaaaaagcagcaaaaagttagcggtta 475  
 QY 433 AGGAAGCTATTTGCTAGGTCGCTGCGTAGTATCATCACGCTC---TGAAAAAGTTTGGAGAGT 489  
 Db 476 aagaagaatacaccagtgagcagcacttctgcacactccgcatcacaatatcgcggaaac 535  
 QY 490 ATATCTCAGAAAGCTATGGAGCGTGTGGGCANCGATGGTGTGATTCACATCGAAGATCTC 549  
 Db 536 tcactgctgacgctatggaaaaagttggttaaaagcgcgtgataccgcttgaggaagcta 595  
 QY 550 GAGGTATGGAACAGAACTTGAAGTGGTGTGAAGCCATTTGACCGTGTGTACTCTGT 609  
 Db 596 agggcaattgaagatgaattgagtgctgtagaaggaatgcaatttgatagaggtactct 655  
 QY 610 CTCATATCATGTGTCAGAGCAATGAAAAATGTTGTCAGACCTTTGAAACCCCATTTATCT 669  
 Db 656 ccccttatttgaacgaacgctgagaaatgaccgctcaattggataatgcttacatcc 715  
 QY 670 TAATCAGCATAAAAAGTCTCAAACTCCAGACATTTTGCACACTACTGTGAGGAAGTTC 729  
 Db 716 tttaacgataaaaaactctagcatgaagacacattcccgctactagaaaaacca 775  
 QY 730 TTAACCAACCCCTCCATTACTCATTTATTCAGATGATGGTGTGAAGCACTTCCAA 789  
 Db 776 tgaaagagggcaaacccgttttaactatcgctgaagacattgagggcggaagctttaacga 835  
 QY 790 CCCTTCTTTGAACAAGATTCGTGGTACTTTCAATTTGTTGCTGTCAAGCCCGCAGGAT 849  
 Db 836 ctctagtggtgaataaataaagggcggtgtgaaatcgagcgggttaaaagctccaggct 895  
 QY 850 TTGGTATGCTGCTTAAGCTATGCTTTGAGACATATGCTATCTTGACAGGTGGTACAGTGA 909  
 Db 896 ttgggacagaagaagaataatgctcaagacatcgctattttaaccggcggtcaagtc 955  
 QY 910 TTACAGAGATCTPAGGATCTGAATTTAAAGATGCTACAATGACAGCCCTTGGACAGGCTG 969  
 Db 956 ttgagcaagaattgggcttgagctctgagaacacgcgtgaagtggagttttttaggcaagctg 1015  
 QY 970 CTAAGATTACAGTTGATAAGTAGACACATAATTTGTAAGGTTTCAGGAAGTTTCAGAA 1029  
 Db 1016 gaaggatttgtagtgcacaaagacaacacccagctcgtagatgcaaaaggccatagcgatg 1075  
 QY 1030 CTATTCTTAACCTATGCTACTGATTAATCGCAATTTAGAACACAACTTCTGACTTTG 1089  
 Db 1076 atgttaagacagagtcgcgcagatcaaaacccaattgcaadtacgacaagcgattatg 1135  
 QY 1090 ACCGTGAAAAAATFACAAGACGTTTGGCGAAATTTAGCTGGTGTGTAGCTGTGTTATCAAG 1149  
 Db 1136 acaaaagaataattgcaagaagaattggctaaactctctggcggtgctggtgattaaaag 1195

QY 1150 TAGGAGCTCCACAGACAGACAGCTTTTAAAGAAATGAAACTTCGATTCAGGATGCTCTAA 1209  
 Db 1196 tggcgctgcgagtggaagtgaatgaaagagaaaaaacccgggtggaatgcgcgttga 1255  
 QY 1210 ATGCTACACGCTGACGCGTTTGAAGAGGTATCGTGTGTTGGTGGAAACAGCACTTATTA 1269  
 Db 1256 gcgcgactaaagcgggttgaaagagggcattgagattggcggtggcggtctctcattc 1315  
 QY 1270 CGGTTTATGAAAAATGACGAGCTCTTTGAGCTTGAAGGCGCATGATGCTGACGCTAA 1329  
 Db 1316 gcgcgctcaaaaagt---gcatttgaattgcagcatgatgaaagagtggtctatgaa 1372  
 QY 1330 TTGTGCTTCTGCTCTAGAGAGCCCTGTAGCTCAAAATGCTTTAAATGCTGGGTACGAAG 1389  
 Db 1373 tcatactgcgcgcatttaaagcccatctagctcaaatcgctatcaacgctgggttatgatg 1432  
 QY 1390 GCTCCGCTAGTATTGCAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATCTCCAA 1449  
 Db 1433 gcgggtggcgtagaagtagaaaaacacagaaaggcatttggttttaaocgctagca 1492  
 QY 1450 CAGGTGAGTGGGTTGATATGATTAAACAGGAATCATTTGACCCCTGTCAAAGTAACACAT 1509  
 Db 1493 atggcaagtagtggtatgtttaaagaaggaatttgaacctttaaaagttagaaagga 1552  
 QY 1510 CAGCGCTTCAAAATGACGCTTCTGTAGTCTTATTTTGAACAACAGAGAGTGTGTTG 1569  
 Db 1553 tcgctctacaaaatgggttctcggttttcaagcctgcttttaaccacagaagccacgt-- 1610  
 QY 1570 CTAAATACCTGAAACAGCTACGCCAGCCAGCAATGCCAGAGGTTATGATCCAGGAA 1629  
 Db 1611 -gcataaatcaagaagaaaaagcactccgcaatgcctgatggtggcgatggcg 1669  
 QY 1630 TGATGGTGGGATGGCGG 1648  
 Db 1670 gtatggagggcatggcg 1688

RESULT 3  
 US-08-466-662-7 ; Sequence 7, Application US/08466662B  
 ; Patent No. 6130059  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Covacci, Antonello  
 ; APPLICANT: Bugnoli, Massimo  
 ; APPLICANT: Telford, John  
 ; APPLICANT: Macchia, Giovanni  
 ; APPLICANT: Rappuoli, Rino  
 ; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And  
 ; FILE REFERENCE: CHIR0057  
 ; CURRENT APPLICATION NUMBER: US/08/466,662B  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: Patent Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 1838  
 ; TYPE: DNA  
 ; ORGANISM: Helicobacter pylori  
 US-08-466-662-7

Query Match 31.8%; Score 528.6; DB 3; Length 1838;  
 Best Local Similarity 58.9%; Pred. No. 1.4e-146;  
 Matches 966; Conservative 0; Mismatches 664; Indels 9; Gaps 3;

QY 13 ATATGCAAAAGAAATTAATTTTCAGCAGATGCGCGTGTGCTCATGTTGCGGAGTTG 72  
 Db 56 aaatggcaaaagaatcaaatcttcagatagtcgagaaaaccccttttatttgaaggcgta 115  
 QY 73 ATATGTTACAGATACCGTCAAAAGTAACTGCTTAAAGGCGCAATTTGTTCTT 132  
 Db 116 ggaactccatgacgctgtcaagtaaccatgggccaagagcaggaatgtattgatcc 175



; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0137-02000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2284 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-467-822-28

Query Match 29.9%; Score 496.6; DB 2; Length 2284;

Best Local Similarity 58.1%; Pred. No. 4.4e-137;

Matches 953; Conservative 0; Mismatches 674; Indels 12; Gaps 4;

```
QY 13 ATATGGCAAGAAATCAAAATTTTCAGCAGATGGCGGTGTCGCATGGTGGCGGAGTTG 72
Db 504 AAATGGCAAGAAATCAAAATTTTCAGATAGCGCAAGAAACCTTTTATTTGAAGCGCTAA 563
QY 73 ATATGTTAGCAGATACCGTCAAAAGTAACGCTGTGCTCTAAAGGCGCAATCTTCTTCTTG 132
Db 564 GACAATCCATGACGTGTCAAAAGTAACCATGGGCGCAAGGAGGAGGAGCGTGTGTATCC 623
QY 133 AAAAAGCTTTTGTCTCTCCCTTAATTAATACGCGGGTAACCATTTGCTTAAAGAGATCG 192
Db 624 AAAAAGCTATGGCGTCCCAAGCATCACCAAGACGCGTGAAGCGTGTAAAGAGATTG 683
QY 193 AATTAGAGATCAATTTGAAAACATGGGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAAA 252
Db 684 AATTAAGTTGCCCGTGGCTTAAGATAGGCGCTCAGCTCGTTAAAGAGATGCGGAGCAAA 743
QY 253 CCAATGATATTGCTGGTGGAGCGACTACTCAACAGTTTTCACACAGCCATTGCTTC 312
Db 744 CCGTGTATGCGCGCGGATGGGACGACCAACGCGGCTGGTGGCTTATAGCATTTTAA 803
QY 313 ATGAAGGACTTAAATAATGACAGCGAGTGTCTAATCCAAATTTGGTATCCGTCGAGGCAATTG 372
Db 804 AAGAGGCTTGAAGATATCACGCTGGGCTTAACCTATTGAAGTGAACAGGCGCATGG 863
QY 373 AAACAGCAACAGCAACAGCTGTGTAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCA 432
Db 864 ATAAAGCGCTGAAGCGATCATTAATGAGCTTAAAGAGCGAGCAAAAAGTGGGCGGTA 923
QY 433 AGGAAGCTATTGCTCAGTTCGCTGCAGTATCATCACGCTC---TGAAGAAAGTTGGAGAGT 489
Db 924 AAGAAGAAATCACCAAGTAGCAGCCATTTCTGCAAACTCCGATCACAATATCGGGAAC 983
QY 490 ATATCTCAGAAGCTATGGAGCGTGTGGGCAACGATGTTGATTAACCATCGAAGAAATCTC 549
Db 984 TCATCGCTGACGCTATGGAAAAAGTGGTAAAGACGCGTGATCACCGTTGAAGAACTA 1043
QY 550 GAGTATGGAACAGAACTTGAAGTGTGTAAGCGATGCAATTTGACCGTGGTGTACCTGT 609
Db 1044 AGGCACTTGAAGATGAATTAGATGTCTAGTAGGCAATGCAATTTGATAGAGGCTACCTCT 1103
QY 610 CTCAATACATGGTACAGCAAAATGAAATAATGGTTGAGACCTTGAAGCCCATTTATCT 669
Db 1104 CCCTTACTTTGTAACCAAGCGCTGAGAAATGACCGCTCAATTTGGATTAACGCTTACATCC 1163
QY 670 TAATACGGGATAAAAAAGTGTCAAACTCAAGACATTTTGGCCACTACTTTGAGGAAGCTTC 729
Db 1164 TTTTAAGGGATAAAAAATCTTAGCATGAAGACATCTCCCGCTACTAGAAAAACCA 1223
QY 730 TTTAAACCAACCGTCCATTTACTCATTTATTCAGATGATGGATGGTGAAGCACTTCCAA 789
Db 1224 TGAAGAGGGCAAAACCGCTTTTATCATCTGCTGAAGACATTTAGGGGCAAGCTTTAACGA 1283
QY 790 CCCTTGTCTTGAACAGATTCGTGGTACTTTCAATGTGGTGTGCTGTCAAAGCGCCAGAT 849
```

```
Db 1284 CTCTAGTGGTGAATAATAATTAAAGAGCGCTGTGTAATATCCGACGCGTTTAAAGCTCCAGGCT 1343
QY 850 TTGCTGATCGTCTGTAAGACATGCTTGAAGACATTTGCTATCTTACACAGGTGGTACAGTGA 909
Db 1344 TTGGGACAGAGAGAAAGAAATGCTCAAGACATCGCTGTTTAAACCGCGGTCAAGTCA 1403
QY 910 TTACAGAGGATCTAGGACTTGAATTTAAAGATGCTCAATGACAGCCCTTTGGACAGGCTG 969
Db 1404 TTACGGAAGAAATTTGGCTTGAGTCTAGAAAACGCTGAAGTG--GAGTTTTTATAGGCAAG 1460
QY 970 CTAAGATTACAGTTGATAAAGATAGCACAGTAATTTGTTGAAGGTTTCAGGAAGTTTCAGAAG 1029
Db 1461 CGAAGATTGATTGACAAAGACACACACGATGCTAGATGGCAAGGCGCATAGCCATG 1520
QY 1030 CTATTGCTAACCGTATTGCACTGATTAAATCGCAATTTAGAAACAACAACCTTCTGACTTTG 1089
Db 1521 ACGTCAAGAGACAGATCGCGCAAAATCAAAACCCCAAAATTTGCAAGCAGCAAGCGATTACG 1580
QY 1090 ACGTGAACAACTACAAGACGTTTGGCGAAATTTAGCTGGTGGTGTAGCTGTTTATCAAAG 1149
Db 1581 ACAAGAAAAAATTTGCAAGAAAGATTTGGCCAAACTCTCTGGCGGTGTGGCTGTGATTAAAG 1640
QY 1150 TAGGAGCTCCAACAGACAGACGCTTTAAAGAAATGAAACTTCGCATTGAGGATGCTCTAA 1209
Db 1641 TGGCGCTGCGAGTGAAGTGAATGAAGAGAAAAAGAACCCGGGTGATGACGCTTGA 1700
QY 1210 ATGTACACGCTGCGCCGCTTGAAGAGGATTCGTTGCTGGTGGTGGAAACAGCACTTATTA 1269
Db 1701 GCGGACTAAAGCGCGGTTGAAGAGGATTTGATTGGGGCGGTGGCGCCCTCATTC 1760
QY 1270 CGGTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGCGGATGATGCTACTGGAGCTAACA 1329
Db 1761 GCGGCGCCCAAAAAGT---GCATTTGAATTTACACGATGATGAAAAAGTGGGCTATGAAA 1817
QY 1330 TTGTGCTTCGTGCTCTAGAGAGCGCTCTACGCTCAAAATTTGCTTTAAATGCTGGGTACGAAG 1389
Db 1818 TCATCATGCGCGCCCATTAAGCCCATTTAGCTCAATTCGCTATCATGTCGCGGTATGATG 1877
QY 1390 GCTCCGCTAGTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGAGATTAAATGCTGCAA 1449
Db 1878 GCGGTGTGGTCTGAATGAAGTAGAAAAACAGGAAGGCGCATTTTGGTTTAAACCTAGCA 1937
QY 1450 CAGTGTAGTGGTGTGATGATTAATAACAGAGATCAATTTGACCCCTGTCAAAGTACACAGAT 1509
Db 1938 ATGCAAGATGTTGGACATGTTTAAAGAAAGCATTTATGACCCCTTAAAGTAGAAAAGGA 1997
QY 1510 CAGCGCTTCAAAATGACGCTTCTGCTAGCTAGCTTTATTTTGAACACAGACAGCACTGTTG 1569
Db 1998 TCGCTTTACAAAATGCGGTTTCGGTTTCAAGCCCTGCTTTTAAACACAGAACCCACCGT-- 2055
QY 1570 CTATAAACCTGAACCCAGCTACGCGCAGCAATGCCAGCAGAGTATGGATCCAGGAA 1629
Db 2056 -GCATGAATCAAGAAAGAAAAAGCGCCCGCAGCAATGCTGATATGGTGGCATGGCG 2114
QY 1630 TGATGGTGGGATGGGCGG 1648
Db 2115 GAATGGGAGCATGGGCGG 2133
```

## RESULT 5

US-08-997-080-159

; Sequence 159, Application US/08997080

; Patent No. 5968524

; GENERAL INFORMATION:

; APPLICANT: WATSON, JAMES D.

; APPLICANT: TAN, PAUL L.J.

; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-

; NUMBER OF SEQUENCES: 194

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Law Offices of Ann W. Speckman

; STREET: 2601 Elliott Avenue, Suite 4185

; CITY: Seattle

; STATE: WA

;; COUNTRY: USA  
;; ZIP: 98121  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/997,080  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sleath, Janet  
;; REGISTRATION NUMBER: 37,007  
;; REFERENCE/DOCKET NUMBER: 11000.1007  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 206-269-0565  
;; TELEFAX: 206-269-0563  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 159:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1626 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-997-080-159

Query Match 26.9%; Score 446.8; DB 2; Length 1626;

Best Local Similarity 54.9%; Pred. No. 1.9e-122;

Matches 880; Conservative 0; Mismatches 722; Indels 0; Gaps 0;

QY 15 ATGCGAAAAGAAATCAAAATTTTCAGCAGATGCGCGTCTGCCATGTTGCGCGGAGTTGAT 74  
DB 1 ATGCGCAAGACAAATTCGGTATACGAGAGAGCGCGCGTGGCTTCAGCGGGCGCTCAAC 60  
QY 75 ATGTTAGCAGATACCTCAAGCTTAAGCTTGTCTTAAAGGCGCAATTTGTTTGA 134  
DB 61 GCCCTCGCAGACCGCTAAAGTGAGTTGGCCCGCAAGGGTCGCAACCTGCTGGAG 120  
QY 135 AAAGCTTTTGGTTCCTTTAATTAATAGCGGGGTAACCATGCTAAAGAGATCGAA 194  
DB 121 AAGAGTGGGGCGCCCGCCAGATCAACACGATGCTGTCTCCATGCCAAGAGATCGAG 180  
QY 195 TTAGAGATCATTTTCAAAACATGGGAGCAAAATTTGTTCTGAAGTGGCTTCTAAACC 254  
DB 181 CTGGAGGACCGGTACGAGAAGATCGCGCTGAGCTGGTCAAGAGAGTCCCAAGAGACC 240  
QY 255 AATGATATTGCTGGTATGGGACGACTACTGCAACAGATTTTGACACAGCCCATTTTCAT 314  
DB 241 GACGAGCTCGCGGCGCGGACCGACCCACCGCCCGCTGCTCAGCTCTGGTTCGC 300  
QY 315 GAAGGACTAAAATGTGACAGAGTGTCTAATCAATTTGGTATCCGTCGAGGCAATGAA 374  
DB 301 GAAGGCTCGCAAGCTCGCAGCGCGGCAACCGCTCGGCTCAAGGCTGGCATCGAG 360  
QY 375 ACAGCAACAGCAACAGCTTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGGCAAG 434  
DB 361 AAGGCTGTGCGAGGCTGTCAACCGAGTGGCTGAAGTGGCGCAAGGAGTCCGAGACCAAG 420  
QY 435 GAAGCTATTGCTCAGTGTGCTCAGTATCATCACGCTCTGNAAGAGTTGGAGAGTATATC 494  
DB 421 GAGCAGATTCTTGCCACCGCGCGGATTTCCGCGCGGACACCCAGATCGCGAGCTCATC 480  
QY 495 TCAGAGCTATGAGCGTGTGGCAACGATGTTGATTAACATCAAGAAATCTCGAGT 554  
DB 481 GCCGAGGCGCATGGAAGTTCGGCAACGAGGAGTGTATCACCGTGGAGAGTCAACACC 540  
QY 555 ATGGAACAGCAATTCGAAGTGGTTGAAGGCATGCAATTTGACCTGGTTTACCTGTCTCA 614

DB 541 TTCGGCCTCGAGCTCGAGCTACCGAGGGTATCGCGTTGCACAAGGGCTACATCTCGGGT 600  
QY 615 TACATGTCACAGACAATGAAAAATGGTTGCAGACCTTGAAAAACCATTTATCTTAATC 674  
DB 601 TACTTGTGACCGACCGCGGAGCGCGTCTTGAGGATCCCTACATCTCTGCTG 660  
QY 675 ACGGATAAAAAAGTGTCAAAACATTCAGACATTTTCCCACTACTTGAAGAGTTCTTAA 734  
DB 661 GTCAAGTCCCAAGGTGTGACCGCTCAAGGATCTGCTCCGCTGTGGAAGGTCTATCCAG 720  
QY 735 ACCAACGCTCCATTACTTACTTATTGAGATGATGATGATGATGATGATGATGATGATG 794  
DB 721 GCGGCAAGCGCGCTGCTGATCATCGCGGAGGCTGAGGGCGAGGCGCTGTCCAGCGTG 780  
QY 795 GTCTTGAACAAGATTCGTGTACTTTCAATGTGTTGTCTCAAGCGCCAGGATTTGGT 854  
DB 781 GTGGTCAACAAGATTCGCGGACCTTCAAGTCCGCTCGCGCTCAAGGCTTCGGGCTTCGGT 840  
QY 855 GATCGTCGTAAGCTATGCTTGAAGACATTTGCTATTTGACAGGTTGTTACAGTATTACA 914  
DB 841 GACCGCGCAAGCGGATGCTGCGAGGACATGGCCATCTCACCGGTGTGAGGTCTGTCAGC 900  
QY 915 GAGGATCTAGGACTTGAATTAAGATGCTACATACAGACCCCTTGGACAGGCTGCTAAG 974  
DB 901 GAAAGATCGGGCTGTCCCTGGAGACCGCGGAGCTCTGCTGCTGGGCGAGGCCCGCAAG 960  
QY 975 ATTACAGTTGATAAGATAGACAGTAATTTGTTGAAGTTTCAGGAAGTTTCAGAAGCTATT 1034  
DB 961 GTGCTGCTCACCNAAGCAGACACCATCTGCGAGGGCTTCGGGCGATTCGGATGCCATC 1020  
QY 1035 GCTAACCTGATTTGCACTGATTAATGCAATTTAGAAACAACAACCTTCTGACTTTTGACCT 1094  
DB 1021 GCGCGCGGGTGGCTCAGATCCGCGCGAGATCGAGAACAGGACTCCGACTACGACCGC 1080  
QY 1095 GAAAACTACAGAGAGCTTTGGGAAATTTAGCTGGTGGTGTAGCTTTTATCAAGTAGGA 1154  
DB 1081 GAGAAGCTCGAGAGCGCTGGCCAAAGCTGGCGCGGCTGTTCCGCTGATCAAGGCCGA 1140  
QY 1155 GCTCAACAGAGACAGCTTTTAAAGAAATGAACTTCGCAATTCAGGATGCTCTAAATGCT 1214  
DB 1141 GCTGCCACCGAGTGGAGCTCAAGGAGCGCAGCAGCCGATCGAGGACCGCTCCGCAAC 1200  
QY 1215 ACAGTGTGACGCGTTGAAGAAGTATGCTTGTGGTGGTGGGAACAGCACTTATACGCTT 1274  
DB 1201 GCGAAGGCTGCGCTCGAAGAGGCGATCGTGGCGGTGGCGGCTGCTCTGCTGCACTCG 1260  
QY 1275 ATTGAAGATGACAGCTCTTGAAGTGGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTG 1334  
DB 1261 GCTCTGCTGCTGAGAGCTCTGAGGCTGAGGCGGAGGCGGCGGCGGCGGCGGCGGCG 1320  
QY 1335 CTTCGTGCTCTAGAAGAGCTGTAGCTCAATTTGCTTTAAATGCTGGGTACGAAGGCTCC 1394  
DB 1321 CCGGTGGCGCTGCTGCTCGCTCAAGCAGATCGCTTCAAGCGCGCTGAGCGCGCGC 1380  
QY 1395 GTAGTTATTGACAAGTTGAAGAACAGCCCTGCAAGAACAGGATTTAAGTCTGCAACAGT 1454  
DB 1381 GTGCTTGGCAGAAGGTGTCCAACCTGCGCGGCTGCGGCGCTCAAGCGCGGCGGCT 1440  
QY 1455 GAGTGGTGTATGATTAAGAACAGGATCATTTGACCTGTCAAGATACAGATCAGG 1514  
DB 1441 GAGTACGAGGAGCTGCTCAAGCGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG 1500  
QY 1515 CTTCAAATGCAAGCTTCTGCTAGCTAGTCTTATTTTGAACAACAAGAGGAGTGTGCTAAT 1574  
DB 1501 CTGCAAGAGCGGCGTCCATCGGCTCTGTTCTTCAACCGGAGGCGCTGCTGCCGAC 1560  
QY 1575 AAACCTGAACAGCTTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1616  
DB 1561 AAGCGGAGAGGCGTCCGACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1602

RESULT 6  
US-08-997-362-159

```

: Sequence 159, Application US/08997362
: Patent No. 5985287
: GENERAL INFORMATION:
: APPLICANT: Tan, Paul
: APPLICANT: HiYama, Jun
: APPLICANT: Visser, Elizabeth
: APPLICANT: Skinner, Margot
: APPLICANT: Scott, Linda
: APPLICANT: Prestidge, Ross
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
: TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
: NUMBER OF SEQUENCES: 194
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Law Offices of Ann W. Speckman
: STREET: 2601 Elliott Avenue, Suite 4185
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98121
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/997,362
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
: FILING DATE: June 12, 1997
: APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
: FILING DATE: August 29, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sleath, Janet
: REGISTRATION NUMBER: 37,007
: REFERENCE/DOCKET NUMBER: 11000.1002c2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-269-0565
: TELEFAX: 206-269-0563
: TELEX:
: INFORMATION FOR SEQ ID NO: 159:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1626 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-997-362-159

Query Match 26.9%; Score 446.8; DB 2: Length 1626;
Best Local Similarity 54.9%; Pred. No. 1.9e-122;
Matches 880; Conservative 0; Mismatches 722; Indels 0; Gaps

QY 15 ATGCCAAAAGAAATCAAAATTTTCAGCAGATGCGCGCTGCCCATGGTGC CGGAGTTGAT 74
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGCCCAAGACAATTGCGTATGACGAAGAGCGCCGCGTCTGAGCGGGGCGCTCAAC 60

QY 75 ATGTTAGCAGATACCGTCAAGTAACGCTTGGTCTTAAGGGGGCAATGTTGTTCTTGA 134
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GCCTTCGCAGACCGCTAAAGGTGACGTTGGGCCCGCAAGGGTCGCAACCTCGTGTG 120

QY 135 AAAGCTTTTGTTCTCCCTTAATTACTAATGAGGGGTAACCATTCGTAAGAGATCG 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AGAAGTGGGGCGCCCGCCAGCATCCACGATGGTGTGTCATGCCAAGGAGATCGAG 180

QY 195 TTAGAAGATCATTTTGAACAACATGGGAGCAAAATTTGGTCTCTCAAGTGGCTCTTAA 254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CTGAGGAGCCCGTACGAGAAGATCGCGCTGAGCTGGTCAAGAGGTGCCAAGAGACC 240

QY 255 AATGATATTGCTGGTGATGGGAGCTACTGCAACAGTTTTTGACAAACGCCATTTGTT 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GACCAAGCTCGCGGGCAGCGACCAACCGCCAGCGTGTGCTCAGGCTCTGTTGCTGCG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 1395 GTAGTTATTGACAAAGTTGAAACACCCCTGACAGGACAGGATTTAATGCTGCAACAGGT 1454  
 DB 1381 GTCTTCCGCGAGAGAGTGTCCAACTGCGCGCGGTGACGGCTCAACGCGCGACCGT 1440  
 QY 1455 GAGTGGTGTATGATTAACACAGCAATCATTTGACCTGTCAACAGTACACAGTACGCG 1514  
 DB 1441 GAGTACAGAGCTGCTCAAGCGCGGTGCGGCGCGGTGAAGGTCAACCGCGTGGCG 1500  
 QY 1515 CTTCAAAATGACGCTTCTGTAGTCTATTTTACACAGCAAGCAGTGTGTGCTAAT 1574  
 DB 1501 CTGACAGACGCGGTGCTGCTGCGGTCTGCTTCAACACCGCGCGTGTGCGCGAC 1560  
 QY 1575 ARACTGACACAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1616  
 DB 1561 AAGCGGAGAGGCGTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1602

RESULT 7  
 US-09-095-855-159  
 ; Sequence 159, Application US/09095855  
 ; Patent No. 6160093  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tan, Paul  
 ; APPLICANT: Visser, Elizabeth  
 ; APPLICANT: Skinner, Margot  
 ; APPLICANT: Prestidge, Ross  
 ; TITLE OF INVENTION: Compounds and Methods for  
 ; TREATMENT: Treatment and Diagnosis of Mycobacterial Infections  
 ; NUMBER OF SEQUENCES: 208  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Law Offices of Ann W. Speckman  
 ; STREET: 2601 Elliott Avenue, Suite 4185  
 ; CITY: Seattle  
 ; STATE: WA  
 ; COUNTRY: USA  
 ; ZIP: 98121

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/095.855  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/705.347  
 FILING DATE: 29-AUG-1996  
 APPLICATION NUMBER: 08/873.970  
 FILING DATE: 12-JUN-1997  
 APPLICATION NUMBER: 08/997.362  
 FILING DATE: 23-DEC-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sleath, Janet  
 REGISTRATION NUMBER: 37,007  
 REFERENCE/DOCKET NUMBER: 11000.1002c3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-269-0565  
 TELEFAX: 206-269-0563  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 159:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1626 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-09-095-855-159  
 Query Match 26.9%; Score 446.8; DB 3; Length 1626;  
 Best Local Similarity 54.9%; Pred. No. 1.9e-122;  
 Matches 880; Conservative 0; Mismatches 722; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAAATCAAAATTTTACAGATGCGCGTGTCCCATGTCGCGAGTTGAT 74  
 DB 1 ATGGCAAAAGAAATTTGCGTATGACGAAGAGCGCGCGTGGCTCGAGCGGCGCTCAAC 60  
 QY 75 ATGTTAGAGATACCGTCAAGTAAAGCTTGGTCTTAAAGGGCGCAATGTTGTTCTTGA 134  
 DB 61 GCGCTCGAGAGCGCGTAAAGGTGACGTTGGGCGCGAAGGGTGCACGCTGCTGCTGGAG 120  
 QY 135 AAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGTAAACCATTCCTTAAAGAGATCGAA 194  
 DB 121 AGAAGTGGGGCG 180  
 QY 195 TTAGAAGATCATTTTGAAGCAATGGAGCAAAATTTGCTGTCTGAAGTGGCTTCTTAAAC 254  
 DB 181 CTGGAGGACCGGTACGAGAAGATCGCGCTGAGTGGTCAAGAGAGTGCACGAAGAGACC 240  
 QY 255 AATGATATTGCTGGTGTGATGGGAGGAGTACTGCAACAGTGTGTCGAGGCAATGTTCA 314  
 DB 241 GACGAGCTCG 300  
 QY 315 GAAGGACTAAAAAATGTGACAGCAGTGTGTAAGCCCTTGAAGCCATTTGCTCGAGGCAATG 374  
 DB 301 GAAGGCTCGGCAACGTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
 QY 375 ACAGCAACAGCAACAGCTGTGTAAGCCCTTGAAGCCATTTGCTCAACCTGTATCTGGCAAG 434  
 DB 361 AAGGCTGTCGAGGCTGTCAACCCAGTGGTGAAGTGGCGCAAGGAGTTCGAGACCAAG 420  
 QY 435 GAAGCTATTGCTCAGTGTGCTGAGTATCATCAGCTCTGAAAAGTTGGAGAGTATATC 494  
 DB 421 GAGCAGATTTTGCACCG 480  
 QY 495 TCAGAAGCTATGAGCGTGTGGCAACAGTGTGTAATTAACCATTCGAAAGTTCGAGT 554  
 DB 481 GCGAGGCGCATGGCAAGTTCGCAACAGGAGTGTATCATCCGTCGAGGAGTGCAGACAC 540  
 QY 555 ATGGAACAGAACTTCAAGTGGTGAAGGCAATTTGACCGTGGTGTACCTGTCTCAA 614  
 DB 541 TTCGGCTGCGAGCTCGAGCTACCGAGGATGCGCTTCGCAAGGCTACATCTCGGCT 600  
 QY 615 TACATGGTCACAGACAAATGAAATGTTGAGACCTTGAAGCCATTTATCTTAATC 674  
 DB 601 TACTTGTGACCG 660  
 QY 675 ACGGATAAAAAAGTGTCAAAACATTCAGCAATTTTGGCCACTACTTGAGGAAGTTCCTTAA 734  
 DB 661 GTCAGCTCAAGGTGTCGACCGCTCAAGGATCTGCTCCCGCTGCTGGAGAGGTCATCAG 720  
 QY 735 ACCAAGCGTCCATTAATTCAGATGATGTTGAGTGGTGAAGCACTTCCAAACCGCTT 794  
 DB 721 GCGGCAAGCGCTGCTGATCATTCGCGAGGAGCTCGAGGCGGAGGCGCTGTCCAGCTG 780  
 QY 795 GTCCTTGAACAGATTCGTTGGTACTTTCAATGTTGTTGCTGTCTAAAGCGCGAGATTTG 854  
 DB 781 GTGGTCAACAGATTCG 840  
 QY 855 GATCGTCTGTAAGCTATGCTTGAAGCAATTTGTAATCTTGAAGGTTGGTACAGTATTA 914  
 DB 841 GACCGCGCGAGGCGATGCTGCGAGGAGTGGCGCTTCCAGCGGTTGGTGTGCTGCTCAG 900  
 QY 915 GAGGATCTAGGACTTGAATTAAGATGCTAATGACAGCGCTTGGAGAGGCTGCTAAG 974  
 DB 901 GAAAGAGTGGGCTGCTGCTGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960  
 QY 975 ATTACAGTTGATAAGATAGACAGTAAATGTTGAAGTTTCAAGAGTTTCAAGAGCTAT 1034  
 DB 961 GTCTGCTGACCAAGGAGAGACCACTGCTGCGAGGCTCGGGCGGATTCGATGCCATC 1020  
 QY 1035 GCTAACCGCTATTGCACTGATTAATTAATGCAATTTAGAAAAACAACTTCTGACTTTTGACCG 1094  
 DB 1021 GCGGCGCGGTTGGCTCAGATCCGCGCGCGAGATCGAGACAGCGACTCCGACTACGACCG 1080  
 QY 1095 GAAAAACTACAAGAACGTTTGGCGAAATTTAGCTGTGTGTAGCTGTTATCAAGTAGGA 1154

Db	1081	GAGA	CTCG	AGAG	CGCCTG	CCCA	AGCTG	CCGCG	CGGTG	TTCG	GGTG	ATCA	AGGCG	CGGA	1140
Qy	1155	GCTC	CAAC	AGAG	CAGAGCT	TTTAA	AGAAAA	TGA	AACCTT	CGAT	TGAG	GATG	CTCT	AAAT	GC
Db	1141	GCTG	CCAC	CGAG	TGGAG	CTCA	AGGAG	CGCA	AGCAC	CGCAT	CGAG	GAC	CGGCT	CCGC	CAAC
Qy	1215	ACAG	TG	CAG	CGT	TGA	GAAG	GTAT	CGTTC	TG	TGGT	GGTGA	CACAG	CAC	T
Db	1201	GCGA	AGGCT	CGCC	GT	CGAA	GAGG	CGCAT	TC	CCGCT	GGCG	GTGG	CT	TC	TC
Qy	1275	ATTG	AAAA	AGT	AG	CAGCT	CTT	CAGCT	TTC	GAGG	CGAT	GAT	CT	CT	CT
Db	1261	GCTC	TG	CGCT	GGAG	CAG	CACT	CG	GGCT	GAC	GGG	CAG	GAGG	CCAC	CG
Qy	1335	CTTC	GT	GC	TCT	AGA	AG	AGC	CTG	TAC	GTCA	AAAT	TGCT	T	T
Db	1321	CGCG	TGG	CGCT	GT	CGCT	CA	AG	CAGAT	CG	CCCTT	CA	ACG	CG	CG
Qy	1395	GTAG	TAT	TG	ACA	AGT	TG	MAA	ACAG	CCCT	CG	AGAC	CAG	ATT	T
Db	1381	GT	CGT	TG	CGG	AGAA	GGT	GT	CCA	CACT	TG	CCG	CGGT	CA	CG
Qy	1455	GAGT	GGT	TG	AT	GAT	TAA	AC	AG	GAAT	CA	TTC	GAC	CTG	T
Db	1441	GAT	C	AG	GAG	CACT	GCT	CA	NGC	CGG	CGT	CG	CGAC	CGGT	TC
Qy	1515	CTTC	AAAA	TG	CAG	CTT	CTG	TAG	CTAG	CTT	AT	T	T	T	T
Db	1501	CTG	CA	AC	AG	CGG	CGT	TCC	AT	CG	CGCT	CT	GT	T	T
Qy	1575	AAAC	CT	G	AAC	AG	CT	TAC	CGC	AGG	CGCC	AC	AT	TGC	AG
Db	1561	AAGC	CGG	AGAA	GGCG	GT	CCG	CAC	CGCG	GGG	CGC	AC	CG	AC	CG

## 8. RESULT

US-08-997-080-113  
Sequence 113, Application US/08997080  
Patent No. 5968524  
GENERAL INFORMATION:  
APPLICANT: WATSON, JAMES D.  
APPLICANT: TAN, PAUL L.J.  
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-  
NUMBER OF SEQUENCES: 194  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,080  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0363  
TELEX:  
INFORMATION FOR SEQ ID NO: 113:

QY 915 GAGATCTAGGACTTGAATTAAGAAGATGCTACAAATGACAGCCCTTGGACAGGCTGTGAAG 974  
Db 901 GAAAGAGTCGGCTGCCCTGGAGAGCCGCGAGCTCTCGTGTGGGCCAGGCGCGCAAG 960  
QY 975 ATTACAGTTGATAAAGATACAGAGTAATTTGAAGTTTCAGGAAGTTTCAGAGGCTATT 1034  
Db 961 GTCTGCTCACCAGGAGAGACCAATCGTCAGAGGCTCGGCGCATCCGATGCGCATC 1020  
QY 1035 GCTAACCGTATTGACTGATTAATAGCAATTAAGAACAACACTTCTGACTTTGACCGT 1094  
Db 1021 GCCGCGGGTGTCTCAGATCCGCGGAGATCGAGACAGGACACTCCGACTACGACCGC 1080  
QY 1095 GAAAACTACAAGACGTTTGGCGAAATAGCTGGTGGTGTAGCTGTTATCAAGATGGA 1154  
Db 1081 GAGAAGCTGAGAGGCGCTTGGCAAGCTGGCGGGGCTTTCGGGTGATCAAGGCGGA 1140  
QY 1155 GCTCCACAGAGACAGCTTTAAAGAAATGAACTTCGCAATTCAGGATGCTCTAAATGCT 1214  
Db 1141 GCTGCCACCGAGGTGAGCTCAAGGAGCGCAACCGCATCGAGGACCGCGTCCGCAAC 1200  
QY 1215 ACAGTGCAGCGCTTGAAGAAGGTATCGTTGCTGGTGGTGAACAGCACTTATTACGGTT 1274  
Db 1201 GCGAAGGCTGCGCTCGAAGAGGCGATCGTCGCGGCTGGCGGCTGCTGCTGCACTG 1260  
QY 1275 ATTGAAGAGTACAGCTCTTGAAGTGAAGGCGGATGATGCTACTGGAGCTAACATTTG 1334  
Db 1261 GCTCTGCGGTGACAGACCTTGGCCCTGAGCGGCGAGGAGCCACCGGTGCCAACATCGT 1320  
QY 1335 CTTGCTGCTCTAGAGAGGCTGACGTCATAATGCTTTAAATGCTGGTACGAAGGCTCC 1394  
Db 1321 CGGCTGGCGTGTGCGCTCGCTCAAGAGATCGCCCTCAACGGCGGCTGGAGCCCGC 1380  
QY 1395 GTAGTTATTGACAAATTTGAAGAAACAGCCCTGCGAGGAAAGGATTTAATGCTCAACAGT 1454  
Db 1381 GTGTTGCGGAGAGGTGTCACACCTGCCGCGGCTCACGGCTCAACGCGGACCGGT 1440  
QY 1455 GAGTGGTGTGATGATTAAGACAGAAATCAATGACCCCTGTCACAACTAACAGATCAGG 1514  
Db 1441 GAGTACGAGGACCTGCTCAAGGCGGCTGCGCGGAGGCTGAAGTCAACCGCTCGGG 1500  
QY 1515 CTTCAAAATGACGCTTCTGTAGTCTGCTTTATTTTACAAACAGCAAGCAGTGTGCTAAT 1574  
Db 1501 CTGCAAGCGCGGCTCCATCGCGGCTCTGTTCTCACCACCGAGCGCGTGTGCGCGAC 1560  
QY 1575 AAACCTGA 1582  
Db 1561 AAGCCGGA 1568

## RESULT 9

US-08-997-362-113  
; Sequence 113, Application US/08997362  
; Patent No. 5985287

## GENERAL INFORMATION:

; APPLICANT: Tan, Paul  
; APPLICANT: Hiyama, Jun  
; APPLICANT: Visser, Elizabeth  
; APPLICANT: Skinner, Margot  
; APPLICANT: Scott, Linda  
; APPLICANT: Prestidge, Ross  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
; NUMBER OF SEQUENCES: 134  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Ann W. Speckman  
; STREET: 2601 Elliott Avenue, Suite 4185  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98121

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,362  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970  
; FILING DATE: June 12, 1997  
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347  
; FILING DATE: August 29, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sleath, Janet  
; REGISTRATION NUMBER: 37,007  
; REFERENCE/DOCKET NUMBER: 11000.1002c2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-269-0565  
; TELEFAX: 206-269-0563  
; TELEX:

## INFORMATION FOR SEQ ID NO: 113:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1569 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; US-08-997-362-113

Query Match 26.8%; Score 444.8; DB 2; Length 1569;

Best Local Similarity 55.2%; Pred. No. 7.5e-122;

Matches 866; Conservative 0; Mismatches 702; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAATCAAAATTTTCAGCAGATGCCGCTGCTGCCATGTCGCGGAGTTGAT 74  
Db 1 ATGGCCAAAGACAAATTCGGTATGCAAGAGAGCCGCGGCTCGAGCGGGCCCTCAAC 60  
QY 75 ATGTTAGCAGATACCTCAAAAGTAAAGCTTGGTCTTAAAGGGCGCAATGTTGTTTGA 134  
Db 61 GCGCTCGCAGACCGCTAAAGGTGAGGTTGGGCCCGCAAGGTCGCAACGTCGTCGTCGAG 120  
QY 135 AAAGCTTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 194  
Db 121 AAGAAGTGGGGCGGCGCCACCATCAACAGATGTTGTGTCCATCCCAAGGAGATCGAG 180  
QY 195 TTAGAAGATCATTTTGAACACATGGAGCAAAATTTGGTGTCTGAAGTGCCTCTAAACC 254  
Db 181 CTGAGGACCCGTACGAGAAGATCGCGCTGAGCTGTTCAAGAGGTCGCCAAGAGACCC 240  
QY 255 AATGATATTGCTGGTATGGGACGACTACTGCAACAGTTTGTGACAAAGCCATTTGTTCA 314  
Db 241 GAGCAGCTCGCGGGGAGCGGACCCACCGCCAGCGTCTCGCTCAGGCTCTGGTTCGC 300  
QY 315 GAAGGACTAAAAATGTGACAGAGGTGCTAATTAATTAATTAATTAATTAATTAATTAAT 374  
Db 301 GAAGGCTCGCAAGCTCGCAGCGCGGCAACCGCTCGGCTCAAGCTGCGCATCGAG 360  
QY 375 ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGCAAG 434  
Db 361 AAGGCTGTGAGGCTGTACCCAGTCGCTGTAAGTTCGCGCAAGGAGGTGCGAGACCAAG 420  
QY 435 GAAGTATTGCTCAGTCTGCTCAGTATCATCAGCTCTGTAAGGAGTTCGAGCATATATC 494  
Db 421 GAGCAGATTTCTGCCACCGCGGCGGATTTCCGCGGGGACACCCAGCATCGCGAGCTATC 480  
QY 495 TCAGAGCTATGAGCGGTGTCGCAACGATGTTGTTGATTAATTAATTAATTAATTAATTAAT 554  
Db 481 GCGGAGGCGCATGCAAGGTCGCAACGAGGAGGTGTCATCACCGTCGAGGAGTCAACACC 540  
QY 555 ATGGAACAGAACTTGAAGTGGTTGAAGGCTGCAATTTGACCGGTGGTTACCTGCTCA 614  
Db 541 TTCGCGCTGACGCTCAGCTCAGCTCAGGAGGTATGCGCTTCGCAAGGAGGTACATCTCGGT 600



```
QY 615 TACATGGTCACAGACATGAAATGGTTGCAGACCTTGAAACCCATTATCTTAATC 674
Db 601 TACTTCGTCCAGCCGCGGAGAGCGGTCCTGAGGATCCCTACATCCTGCTG 660
QY 675 ACGGATAAAAAGTGTCAACATCCAGACATTTGCGCACTACTTGAGGAAGTCTTAAA 734
Db 661 GTCAAGTCCAAGGTGCGACCGTCAGGATCTGTCCTCGCTGCTGGAAGAGGTCAATCCAG 720
QY 735 ACCAACCGTCCATTACTTATTTAGAGATGATCTGGATGCTGAAGACATTTCCAACCCCTT 794
Db 721 GCCGGCAAGCGCTGCTGATCATCGCGAGGAGCTGCGAGGCGAGGCCCTGTCCACGCTG 780
QY 795 GTCTTGAAAGATTCGTGTACTTCAATGTGTGCTGTCAAGGCGCCAGGATTTGGT 854
Db 781 GTGTCAACAAGATCCGCGCACCTTCAAGTCGCTGCCCTCAAGGCTCCGGGCTTCGGT 840
QY 855 GATCGTCGTAAGCTATGCTTGAAGACATGCTATCTTGACAGTGTGTACAGTATTAACA 914
Db 841 GACCGCGCAAGGAGTGTGTCAGGACATGCCATCTCACCGGTGTCAGGTCGTCAGC 900
QY 915 GAGGATCTAGGACTTGAATTTAAAGATGCTACAAATGACAGCCCTTGGACAGGCTGCTAAG 974
Db 901 GAAAGAGTCGGGTGTCCTCGGAGACCGCGAGCTCTCGCTGCTGGGCCAGGCCGCAAG 960
QY 975 ATTACAGTTGATTAAGATACACAGTAATTTGTTGAAGGTTTCAGAGGTTTCAGAGCTATT 1034
Db 961 GTGTGTCGTACCAAGGACGAGACACCATCTGTCAGGCTCGGCGCATTCGGATGCCATC 1020
QY 1035 GCTAACCGTATTCACCTGATTAAATCGCAATTAGAACACAACTCTGACTTTGACCGT 1094
Db 1021 GCCGGCGGTGCTCGATCCGCGCGGAGATCGAGAACAGGACTCCGACTACGACCGC 1080
QY 1095 GAAAACTACAAGAACCTTTGGCGAAATAGCTGGTGTAGCTGTATCAAAAGTAGGA 1154
Db 1081 GAGAAGCTGACGAGCGCTTGCGCAAGCTGCGCGCGGTGTGGGTGATCAAGGCCGA 1140
QY 1155 GCTCCACAGACAGACGCTTTAAAGAAATGAATCTGCGATTCAGGATGCTCTAAATGCT 1214
Db 1141 GCTGCCACCGGTTGGAGCTCAAGGAGCGCAAGCACCGCATCGAGGACGCGCTCCGCAAC 1200
QY 1215 ACAGTGCACCGCTGAAGAGGTATGTTGCTGGTGTGAACAGCACTTATACGGTT 1274
Db 1201 GCGAAGCTCGCGTTCGAGAGGCTATGTCGCGGTGGCGGCGGTGCTCTGTCGAGTCG 1260
QY 1275 ATTGAAAAAGTAGCAGCTCTGAGCTTGAGGCGCATGATCTACTTGACCGTAACATTTG 1334
Db 1261 GCTCCTCGCTGGAGACCTCGGCCTCACGCGGCGAGAGCCACCGGTGCAACATGTC 1320
QY 1335 CTTCGTGCTCTAGAGAGCTGTACGTCAAAATGCTTTAAATGCTGGGTACGAGGCTCC 1394
Db 1321 CGCGTGGCGTGTGCGCTCGGCTCAAGCAGATCGCCTTCAACGCGCGCTTGGAGCCCGC 1380
QY 1395 GTAGTTATTACAAAGTTGAAAAACAGCCCTGCAGGAACAGATTTATGCTGCAACAGGT 1454
Db 1381 GTGTTGCCGAGAGGTGTCCACCTGTCGCGCGGCTCACGGCTCAACGCGCGGACCGGT 1440
QY 1455 GAGTGGTGTGATGATTAAGAGGAATCATTCACCTGTCAAAGTAAACAGCATCAGCG 1514
Db 1441 GAGTACGAGGACCTGCTCAAGCGCGGCTGCGCGACCGGTGAAGGTCAACCGCTCGCG 1500
QY 1515 CTTCAAAATCGAGCTCTGTAGTGTCTTATTTGACACAGAGAGTGTGTGTTAAT 1574
Db 1501 CTGCAGAACCGCGTCCATCGGGTCTCTTCTCTACCAACCGAGGCGGTCTGTCGCGGAC 1560
QY 1575 AAACCTGA 1582
Db 1561 AAGCCGA 1568
```

RESULT 10

US-09-095-855-113

; Sequence 113, Application US/09095855

; Patent No. 616093

```
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 1569 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-095-855-113
```

```
Query Match 26.8%; Score 444.8; DB 3; Length 1569;
Best Local Similarity 55.2%; Pred. No. 7.5e-122;
Matches 866; Conservative 0; Mismatches 702; Indels 0; Gaps 0;

QY 15 ATGGCAAGAAATCAAAATTTTCAGCAGATGCGGCTGCTGCCATGGTGGCGGAGTTGAT 74
Db 1 ATGGCAAGAAATTTGCTATGACGAAGAGCGCGCTCGAGCGGGGCTCAAC 60
QY 75 ATGTACGATACCGTCAAGTAACGCTTGGTCTTAAAGGCGCAATGTTGTTCTTGA 134
Db 61 GCCTTCGACAGCGCGTAAAGGTGACGTTGGGCCGGAAGGTCGCAACGCTGCTGGAG 120
QY 135 AAAGCTTTTGGTCTCCCTTAATTAATGAGGGGTAACCATTTGCTAAAGAGATCGAA 194
Db 121 AAGAAGTGGGCGCCCCCAGCATCACCAACGATGGTGTCCATCGCAAGGATCGAG 180
QY 195 TTAGAAGATCATTTGAAACATGGGAGCAAAATTTGCTGTGAAGTGGCTTCTAAAC 254
Db 181 CTGAGGAGCCGTACGAGAAGATCGGCGCTGAGCTGCTCAAGAGAGTTCGCAAGAGACC 240
QY 255 AATGATATTGCTGCTGATGGAGCGACTACTGCAACAGCTTTTGACACAGCCATTGTT 314
Db 241 GACGAGCTCGGGGCGGACGCGCACCAACCGCCACCGCTGCTCGCTCAGGCTCTGGTTCG 300
```



Db 1 ATGCGAAGATTCTGAAGTTCGACGAGACGCCCGTTCGCCCTTGAGCGCGCGTGAAC 60  
Qy 75 ATGTTAGAGATACCGTCAAGATTAACGTTGGTCTCTAAAGGCGCAATGTTGTTCTTGA 134  
Db 61 CAGCTGGCCGACACCGTCAAGTGACCATCGGCCCAAGGCCGCAACGCTGTCATCGAC 120  
Qy 135 AAAGCTTTTGGTCTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194  
Db 121 AAGAAGTTGGCGCCCGACCATCACCACGACGCGCTCACCATCGCCCGTGGAGTTCGAG 180  
Qy 195 TTAGAAGATCATTTTGAAGACATGGGAGCAAAATGTTGTTGTTGTTGTTGTTGTTGTTG 254  
Db 181 TGGACGACCGGTACGAGAACCTCGCGCCGACGCTCGTCAAGGAGGTGGCCACCAAGACC 240  
Qy 255 AATGATATTGCTGTGATGGAGACGACTACTGCAACAGTTTGTGACAAAGCAATGTTTCA 314  
Db 241 AACGACATCGCGGTGACGCGACCAACCGCGGACCGTGTGCGCCGAGGCGCTGTCGCG 300  
Qy 315 GAAGGACTAAATAATGTGACGAGGTGCTTAATCCCAATTTGTTATCCGTCGAGGCAATGA 374  
Db 301 GAGGCGCTGGCAACGTCGCGCGCGCGCGCTTCCCGCGCGCTTGAAGAGGGCATCGAC 360  
Qy 375 ACAGCAACAGCAGCATGTTGAAGCCCTTGAAGCCATTTGCTCAACCTGTATCTGGCAAG 434  
Db 361 GCGCGCTGCGCGCGCTGTCGCGCGAGCTGCTCGACACCGCGCGCGCGATCGAGCAAG 420  
Qy 435 GAAGCTATTGCTGAGTGGTGTGATGATCATCATCATCATCATCATCATCATCATCATCAT 494  
Db 421 TCCGACATCGCGCGCTGCGCGCGCTGTCGCGCGAGCAAGCAAGTGTGCGCGAGCTCATC 480  
Qy 495 TCAGAACTATTGAGCGTGTGGGCAACGATGGTGTGATTTACCATCGAAGAAATCTCGAGGT 554  
Db 481 GCGGAGCGATGGAACAGGTGCGCGAGGCGGTGTCTACCGCTCGAGAGTCCAAACACC 540  
Qy 555 ATGGAACAGACATGTAAGTGGTTGAAGCATGCAATTTGACCGTGTACTGCTCTCAA 614  
Db 541 TTCGGTCTCACTGGACTTCCAGGAGGATGCGCTTCACAAAGGCTACCTGTCGCCG 600  
Qy 615 TACATGTCTACAGCAATGAATAATGGTTGACACCTTGAACACCATTTATCTTAATC 674  
Db 601 TACATGTGACGACACGAGGAGGTGAGGCGCTGCTCGACGACCGGTACATCTGATC 660  
Qy 675 ACGATAAAAGTGTCAACATCCACAGATTTTCCACTACTTGTAGGAAAGTCTTAA 734  
Db 661 CACGAGGCAAGATCGGTTGATCCAGGACCTGCTGCGCTGTGGAAGGTCATCCAG 720  
Qy 735 AC-----CAACCGTCCATTAATTTGACATGATGTTGATGTTGAAGCACTTCCA 788  
Db 721 GCGGTGCTTCAAGCGCTGCTGATCATCGCGAGGACGTCGAGGCGGCGCTGTG 780  
Qy 789 ACCCTGCTTGAACAGATTCGTTGTTGTTTCAATGTTGTTGTTGTTGTTGTTGTTGTTG 848  
Db 781 ACCCTGTTGTTCAACAGATTCGCGGACGCTTCAACCGCTGCGCGTCAAGGCGCGCG 840  
Qy 849 TTTGCTGATCGTCAAGATGCTTCAAGACATTCCTATCTTGAAGGTTGTTGTTGTTGTTG 908  
Db 841 TTTGCTGATCGTCAAGATGCTTCAAGACATTCCTATCTTGAAGGTTGTTGTTGTTGTTG 900  
Qy 909 ATTACAGGATCTAGGATTTGAATTAATAAGATGCTACATGACAGCCCTTGGACAGCT 968  
Db 901 ATCCCGGAGGAGTTCGCTCAAGCTGACCGCGCGCTGTCGAGCTGTCGAGGCGCGCC 960  
Qy 969 GCTAAGATTACAGTTGATAAGATPAGCAGATTAATTTGTAAGGTTTCAGGAAGTTCA 1028  
Db 961 CGCGCGTCACTCCGTTCAACAGGAGGACGACGACCATCGTGGAGCGCGCGCAACGCG 1020  
Qy 1029 GCTATTCTTACCGTATTGACATGATTAATTCGCAATTTAGAACAAACAACTTCTGACT 1088  
Db 1021 GAGTTCAGGCGCGCTGCGCGGATCAAGCGCGGATGAGTGCAGCTCGGACTG 1080  
Qy 1089 GAGCGTGAAGAACTACAGAACGTTTGGCGGAATTAATGCTGTTGTTGTTGTTGTTGTT 1148  
Db 1081 GAGCGGAGAGCTTCCAGGAGCGCTTCCCAAGCTGGCGCGCGCTGTCGTTGATCCG 1140

Qy 1149 GTAGAGCTCAACAGACAGACAGCTTTTAAAGAAATGAACATTCGCATTTAGGATGCTCTA 1208  
Db 1141 GTCGCGCGCGCCAGGCTGAGCTGAAGGAGCGCAACACCGCTCTGGAGGAGCGCATC 1200  
Qy 1209 AATCCTACAGCTGAGCGGTTGAAGAGGTATCTGTTGTTGTTGTTGTTGTTGTTGTTG 1267  
Db 1201 TCCGCGACCGCGCGCGGTCGAGGAGGCGATCTCTCCGTTGTTGTTGTTGTTGTTGTT 1260  
Qy 1268 --TAGGCTATTGAAAAGTAGCAGCTTCTGAGCTTGAAGGCGGATGATGCTTACTTGGACGT 1325  
Db 1261 CACGCGCTCAAGTTCCTGGAGCAACCTCGCGCGCACCGCGGAGCGGCGCGGTTGTC 1320  
Qy 1326 AACATTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1385  
Db 1321 GCGTCTGTCGCGCGCGCGGTCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
Qy 1386 GAAGGCTCCGCTAGTTATTGACAAGTTGAAAACAGCCCTGCGAGAACAGGATTTAATGCT 1445  
Db 1381 GAGGCTACGCTCATCACCAACAGGTGGCGAGCTGCACAAGGCGCGGCTTCAACCG 1440  
Qy 1446 GCAACAGGTGAGTGGTGTATGATTAATAACAGGAATCATTTGACCTGTCAAAAGTACA 1505  
Db 1441 GCCACCGGAGTACGCGGACCTGTTCAAGCGCGGCTCATCGACCGGTCAGGTCACC 1500  
Qy 1506 CGATCAGCGCTTCAAAATGACAGTCTCTGTAGCTAGTCTTATTTTGAACAAAGACAGTT 1565  
Db 1501 GCGTCCGCTGAGAGACGCGGCTCCTCATCGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 1560  
Qy 1566 GTTCTTAATAACCTGAACAGCTACCGCAGCGCAG 1602  
Db 1561 GTCGTCGAGAAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1597  
  
RESULT 12  
US-09-031-606-10  
; Sequence 10, Application US/09031606  
; Patent No. 6153404  
; GENERAL INFORMATION:  
; APPLICANT: MAZODIER, Philippe  
; APPLICANT: GUGLIEMI, Gerard  
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE  
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: George Mason Bldg., Washington & Prince Sts.  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031,606  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/050,313  
; FILING DATE: 10-MAY-1993  
; APPLICATION NUMBER: -FR 9011186  
; FILING DATE: 10-SEP-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crane-Feury, Sharon E  
; REGISTRATION NUMBER: 36,113  
; REFERENCE/DOCKET NUMBER: 010830-035  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 10:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
US-09-031-606-10

```

Query Match 25.1%; Score 416.6; DB 3; Length 1620;  
 Best Local Similarity 54.7%; Pred. No. 1.7e-113;  
 Matches 874; Conservative 0; Mismatches 714; Indels 9; Gaps 2;

```

QY 15 ATGGCAAAAGAAATCAATTTTTCAGCAGATGCGCGTGGCTGCGCATGGTGGCGGAGTTGAT 74
DB 1 ATGGCGAAGATTCTGAAGTTTCAGCAGGAGGCGCGCGTGGCGCGCTTGAGCGCGCGGTGAAC 60
QY 75 ATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTTAAAGGGCGCAATGTTGTTCTTAA 134
DB 61 CAGCTGGCGCACACCGTCAAGGTGACCATCGGCCCAAGGGCGCAACGTGCTCATCGAC 120
QY 135 AAGCTTTTGGTCTCCCTTAATTAATGATGAGGGGTAAACATTGCTTAAGAGATCGAA 194
DB 121 AAGAAGTTTCGGGCGCCCGACCATCAACAGACGGCGTCAACATCGCCCGTGAAGTCCGAG 180
QY 195 TTAGAAGATCAATTTGAACAACTGGGAGCAAAATTTGGTCTCTGAAGTGGCTTCTAAAC 254
DB 181 TCGGACGACCGGTACGAGAACTTCGGCGCCAGCTGCTCAAGGAGGTGGCGACCAAGACC 240
QY 255 AATGATTAATTTGGTGGTGGGAGCTACTGCAACAGATTTTGAACAAAGCCATTTGTTAT 314
DB 241 AAGCATATCGCGGTGACGCGACCAACCGCGACGCTGCTGCGCCAGCGCTGGTCCGC 300
QY 315 GAAGGACTAAATAATGTGACAGAGGTGCTATCAATTTGATGCTGCGTGGAGGATTTAA 374
DB 301 GAGGGCTTCGCAACGCTCGCGCGCGCTCCCGCGCGCTCAAGAAAGGCGATCGAC 360
QY 375 ACAGCAACAGCAACAGCTGTTGAAGCTTTGAAGCAATTTGCTCAACCTGATCTGCGAAG 434
DB 361 GCGCGCTGCGCGCGCTTCGCGCGAGCTGCTGACACCGCGCGCGGATCGACGACAG 420
QY 435 GAAGCTATTGCTCAGTGCCTGCTGATCATCATCAGCTCTTGAAAGAGTTGGAGAGTATATC 494
DB 421 TCCGACATCGCGCGCTGCGCGCTCTCCGCGAGGACAAAGCAGTGGCGGAGCTCATC 480
QY 495 TCAGAACTATGAGGCTGGGCAACGATGCTGATTTACCATCGAAGATCTCGAGGT 554
DB 481 GCCGAGGCGATGGCAAGGTTCGGCAGGAGCGGTGCTATCAGCGTGGAGGAGTCCAAAC 540
QY 555 ATGGAAACAGAACTTCAAGTGGTTGAAGGCAATTTGACCGTGGTACCTGCTCTCAA 614
DB 541 TTCGGTGTGACCTTGACTTCAACGAGGAGTGGCGCTCGCAAGGGCTTACCTGTCGCCG 600
QY 615 TACATGGTGCACAGCAATGAAAGATGGTTGCGAGACCTTTGAAAGCCATTTATCTTAATC 674
DB 601 TACATGGTGACCGACGAGGCTATGGAGGCGCTCTCGACGACCGGTACATCTGATC 660
QY 675 ACGGATAAAGAGTGTCAACATCAAGACATTTTGGCACTACTTGAAGAGTCTTTAA 734
DB 661 CACCAGGGCAAGATCGGTTGATCGATCGAGGACCTTGTGCGGCTGCTGGAGAGTCAATCC 720
QY 735 AC-----CAACCGTCCATTTACTATTTGAGATGATGTTGATGCTCAACACTTCCA 788
DB 721 GCGGTGGTCTCAACGCGCTGCTGATCATGCGCGGAGGAGTGGAGGGGAGGCGCTGCG 780
QY 789 ACCCTTGTCTTGAACAAGATTCGTGCTACTTTCAATTTGTTGCTGTCTAAAGCGCCAGGA 848
DB 781 ACCCTGGTGTCAACAAGATTCGCGGACGCTTCAACGCGCTGCGGCTCAAGCGCGCCG 840
QY 849 TTTGGTGTGCTGCGTAAAGCTAATGCTTGAAGACATTTGCTATCTTGACAGGTGGTACAGTG 908

```

```

DB 841 TTCCGTGACCGCGCAAGCGGATGCTCGCGGACATGCCACCTCACCAGTGCACCGTC 900
QY 909 ATTACAGAGGATCTAGGACTTGAATTAAGAGATGCTACAATGACACCCCTTGACAGGCT 968
DB 901 ATCGCGGAGGAGGTGCGGCTCAAGCTCGACAGCGCGGCTTGACGCTGCGGACCGCC 960
QY 969 GCTAAGATTACAGTTGATAAAGATAGCACAGTAATTTGTTGAAGTTTCAGAAAGTTTCA 1028
DB 961 CGCGCGCTCACCGTCAACCAAGGACGACACCATCGTGACGCGCGCGCAACGCGGAG 1020
QY 1029 GCTATTGCTAACCGTATTCAGTATTAATCGCAATTAAGAAACAACTTCTGACTTT 1088
DB 1021 GACGTCAGGCGCGCTCGCCAGATCAAGGCGGAGATCGAGTCGACCGACTCGGACTGG 1080
QY 1089 GACCGTGAAGAACTACAACAAGCTTTGGCGAAATTAAGTTCGTTGGTGTAGCTCTTATCAA 1148
DB 1081 GACCGGGAAGCTCCAGAGCGCTTCGCAAGCTGGCGGCGGCTGCGGTGATCCGC 1140
QY 1149 GTAGGAGCTCCAAACAGAGACAGCTTTTAAAGAAATGAAACTTCGCAATTGAGGATGCTTA 1208
DB 1141 GTCGCGCGCGCACCGAGGTCGAGCTGAAGGAGCGCAACCGCTCTGAGGACGCGCATC 1200
QY 1209 AATGCTACAGCTGACGCGCTTGAAGAGTATGCTGTTGGTGGTGGAAACAGCACTTAT- 1267
DB 1201 TCCGCGACCGCGCGGCTCGAGGAGGAGCATGCTCTCGGTTGGTCCGCGCTGGTC 1260
QY 1268 --TAGGTTATTTGAAAAAGTACAGCTCTTGAAGTTGAGGCGGATGATGCTACTGACCT 1325
DB 1261 CAGCGCGTCAAGGTCCTGAGCAGCAACCTTCGCGCGGACCGCGGAGGCGGCGGCTGTC 1320
QY 1326 AACATTGCTTCGTCTCTAGAAGAGCGCTGTACGTCAAATTTGCTTTAATGCTGGGTAC 1385
DB 1321 GCGGTGCTCGCGCGCGCGCTCGAGCGCTGCGGTGATCGCGGAGAACCGCGGCGCTC 1380
QY 1386 GAAGGTCGCTAGTATTGACAAAGTTGAAAACAGCCCTGACGAGAACAGGATTTAATGCT 1445
DB 1381 GAGGCTAGCTCATCACCACCAAGGTGGCGGAGCTCGACAAGGCGGCGGCTTCAACGG 1440
QY 1446 GCAACAGGTGAGTGGTTGATATGATTAACACAGAAATCATTGACCTCTCAAGTAACA 1505
DB 1441 GCGCGGAGTACGCGGACCTGTCAGGCGGCGGTCTGACCCCGTCAAGGTCAAC 1500
QY 1506 CGATCAGCGCTCAAAATGACGCTTCTGTAGTACTTATTATTGACAAACAGAGCAGTT 1565
DB 1501 GCGTCCGCTTGAGAACGCGCGCTCCATCGCTCTCTCTGAGACCGGAGACCGCTG 1560
QY 1566 GTTGTATTAACCTGAGACCGCTAGCGCGGCGGAG 1602
DB 1561 GTCGTGAGAGCGCGCGGAGGAGGAGCGCGGAGCGG 1597

```

```

RESULT 13
US-08-461-775-11
; Sequence 11, Application US/08461775
; Patent No. 5858773
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,775  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/050,313  
FILING DATE: 10-MAY-1993  
APPLICATION NUMBER: FR 9011186  
FILING DATE: 10-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
TELEPHONE/DOCKET NUMBER: 010830-035  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2668 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-461-775-11

Query Match 24.7%; Score 411; DB 2: Length 2668;  
Best Local Similarity 54.5%; Pred. No. 9.9e-112;

Matches 872; Conservative 0; Mismatches 720; Indels 9; Gaps 2;

```
QY 11 TCATATGGCAAAAGAAATCAAAATTTTCAGCAGATGGCGCTGCTGCCATGGTGGCGGAGT 70
   || || || || || || || || || || || || || || || || || || || || || ||
Db 845 TCCATATGGCAAGATTCGAAGTTTCAGCAGAGGACGCCGCTCGCGCCCTTGAGCGCGCGT 904

QY 71 TGATATGTAGCAGATACCGTCAAGTACGCTTGTCTTAAAGGGGGGCAATGTTGTTCT 130
   || || || || || || || || || || || || || || || || || || || || || ||
Db 905 GAACACAGCTGGCGACACCGTCAAGTGTACCATCGGCCCCCAAGGGCGGCAACGTCGTCT 964

QY 131 TGAAGAAGCTTTTGGTCTCCCTTAATTAATGACGGGTAAACCATTCCTTAAAGAGAT 190
   || || || || || || || || || || || || || || || || || || || || || ||
Db 965 CGACAGAAGTTGGCGCCCGCCGACCATCACCACGAGCGGCTCACCATCGCCGCTGAGGT 1024

QY 191 CGAATTAGAGATCAATTTTGAAGCAATGGGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAA 250
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1025 CGAGTGGCAGACCGGTACGAGAACCTCGCGCCCGAGCTCGTCAAGGAGGTGGCGACCAA 1084

QY 251 AACCAATGATTTGCTGGTGTAGGAGGACTACTGCAACAGTTTTGACACAAAGCCATTGT 310
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1085 GACCAACGACATCGCGGTGACGGCACCCACCGCGGACCGTCTGCGCCGAGCGCTGTT 1144

QY 311 TCATGAAGGACTAAAAATGTGACAGCAGGTGCTAATCCAAATTTGGTATCGTTCGAGGCA 370
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1145 CCAGGAGGCTCGGCAACGTCGCGCGCGCGCTCGCGCGCGCTCGCGCGCGCTGAGAGGCA 1204

QY 371 TGAACAGCAACAGCAACAGCTTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGG 430
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1205 CGACGCGCGCTCGCGCGCTCTCCCGAGCTGCTCGACACCGCGCGCGCGATCGACGA 1264

QY 431 CAAGGAGCTATTGCTCAGTGTCTGAGTATCATCAGCTCTGAAAAAGTTGAGAGTA 490
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1265 CAAGTCCGACATCGCGCGCTCGCGCGCTCTCCCGGAGGCAAGCAGGTTCGCGGAGCT 1324

QY 491 TATCTCAGAGCTATGAGGCTGTGGCAACGATGTTGATTAACCATCGAAGATCTCG 550
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1325 CATCGCGGAGGCGATGACNAGTGGCAGGACCGGTGTCATCAAGCTCGAGGAGTCAA 1384

QY 551 AGGTATGAAGCAACAACTTGAAGTGTGTGAAGGCTGCAATTTGACCGTGGTTACCTGTC 610
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1385 CACCTTCGGTGTGACCTTGACTTCACCGAGGCGCTTCGACACAGGCTACCTGTC 1444

QY 611 TCAATACATGGTCAACAGCAATGAAAAATGTTTGGCAGACCTTTGAAACCCATTATCTT 670
   || || || || || || || || || || || || || || || || || || || || || ||
```

## RESULT 14

US-09-031-606-11  
; Sequence 11, Application US/09031606  
; Patent No. 6153404  
; GENERAL INFORMATION:  
; APPLICANT: MAZODIER, Philippe

```
Db 1445 CCGTACATGGTGACCGACAGGAGCGGTATGGAGCGCGTCTCTGACGACCCGTACATCCT 1504
QY 671 AATCAGGATAAAAAAGTGTCAAAATCCAAAGACATTTTCCCACTACTTTCAGGAAGTTCT 730
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1505 GATCCACCGGCAAGATCGGTTTCGATCCAGGACCTGCTCCCGCTGCTGGAGAAGTCTAT 1564
QY 731 TAAAC-----CAACCGTCCATTACTTATTTGACAGATGATGTGGATGGTGAAGCACT 784
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1565 CAGGCGGCTGGTCCAAAGCGCTGCTGATCATCGCGAGGAGCTCGAGGCGGAGGCCCT 1624
QY 785 TCCAAACCTTCTCTTGAACAAGATTCGTGTACTTTCAATGTGGTTCCTCTCAAGCGCC 844
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1625 GTCGACCTTGGTGTCAACAAAGATTCGCGGCACGTTCAACGCGCTGCGCCGCAAGCGCC 1684
QY 845 AGGATTTGGTGCATCGTAAAGCTATCTTGAAGACATTTGCTATCTTGACAGTGTGATC 904
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1685 CGGCTTCGGTGACCGCGCAAGCGGATGCTCGCGACATGCCACCTCACCGGTGCCAC 1744
QY 905 AGTGATTACAGAGATCTAGGACTTGAATTTAAAGATGCTTACAATGACACCCCTTGGACA 964
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1745 CGTCATCGCGGAGGCTCGGCCCTCAAGCTCGACAGCGCGTCTGGACGCTGTGGGCAC 1804
QY 965 GGCTGCTAAGATTACAGTTGATAAAGATACACAGTAAATTTGTTGAAGTTTCAAGATTC 1024
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1805 CGCCCGCGCGTCAACGATCACCAGGACGACGACCATCTGAGACCTGGAGAAGGACGC 1864
QY 1025 AGAAGCTATTGCTAACCGTATTCACCTGATTAAATCCAAATAGAAAACAACCTTCTGA 1084
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1865 CGAGGAGTCCAGGCGCGCTCGCCAGATCAAGGCGGAGATCGATGTCGACCGACTCGGA 1924
QY 1085 CTTTGACCGGTGAAAAACTACAAAGACGTTTGGGAAATTAGCTGGTGGTGTAGCTGTAT 1144
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1925 CTGGACCGGAGAGTCCAGGAGCGCTCGCAAGCTGCGCGCGCGCTCTCGCTGAT 1984
QY 1145 CAAAGTAGGAGCTCCACACAGACAGCGTTTAAAGAAATGAAACTTCGATTTGAGGATGC 1204
   || || || || || || || || || || || || || || || || || || || || || ||
Db 1985 CCAGTGGCGCGCGCCACCGAGTTCGAGCTGAAGAGGCGCAAGCAGCTCTGGAGGACGC 2044
QY 1205 TCTAAATGCTTACAGCTGCAGCGCTTGAAGAGGTATCGTTGCTGGTGGTGAACAGCACT 1264
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2045 CATCTCGCGACCGCGCGCGCTCGAGGAGGAGTCTCTCGGTGGTGGTCCGCGCT 2104
QY 1265 TAT---TACGGTTATTGAAAAAGTAGCAGCTTTGAGCTTTGAGGCGGATGATGCTACTGG 1321
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2105 GGTCCACCGCTCAAGGTCTGGACGACAACTCGCGCGCACCGCGCAGCGGCGCACCGG 2164
QY 1322 ACCTAATTTGCTGCTGCTCTAGAGAGCCTGTACGTCAAAATGCTTTAATGCTGG 1381
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2165 TGTGCGGCTGCTCGCGCGCGCGCTCGAGCGCTGCGCTGGATCGCGGAGACGCGG 2224
QY 1382 GTACGAGGCTCGGTAGTTATTGACAAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAA 1441
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2225 CTTGAGGCTAGTCTATCACCACCAAGGTGGCGGAGCTCGAAGGCGCGGCTTCAA 2284
QY 1442 TGCTGAACAGGTGAGTGGTGTGATGATTAAACAGGAATCATTTGACCCCTGCTCAAGT 1501
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2285 CGCGGCCACCGCGAGTACCGGACCTTGTCAAGCGCGCGCTCATCGACCCGCTCAAGT 2344
QY 1502 AACACATCAGCGCTTCAAAATGACGCTTCTGTAGCTAGTCTTATTTTGAACAGAGC 1561
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2345 CACCGCTCGCGCTGGAGAACGCGGCTCATCGCTCCCTGCTCTGACGACCGGAGC 2404
QY 1562 AGTTGTTGCTTAATTAACCTGAACCACTAGCCGAGCGCCAG 1602
   || || || || || || || || || || || || || || || || || || || || || ||
Db 2405 CTTGGTCTCGAAGCGCGCGCGGAGGAGGAGCGCGCGGCGC 2445
```

APPLICANT: GUGLIEMI, Gerard  
 TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE  
 INITIATION OF TRANSCRIPTION  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: George Mason Bldg., Washington & Prince Sts.  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/031,606  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/050,313  
 FILING DATE: 10-MAY-1993  
 APPLICATION NUMBER: FR 9011186  
 FILING DATE: 10-SEP-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 010830-035  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703): 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2668 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-09-031-606-11

Query Match 24.7%; Score 411; DB 3; Length 2668;  
 Best Local Similarity 54.5%; Pred. No. 9.9e-112;  
 Matches 872; Conservative 0; Mismatches 720; Indels 9; Gaps 2;  
 QY 11 TCATATGCGAAAGAAATCAAAATTTTCAGCAGATCGCGTGTGCTGCGCGGAGT 70  
 DB 845 TCCCATGGCGAAGATTCGAAGTTCGACGAGGACCGCCGTCGCGCCCTTGACGCGCGGT 904  
 QY 71 TGATATGTTAGCAGATACCGTCAAAAGTAACGCTTGGTCTTAAAGGCGCAATGTTGTCT 130  
 DB 905 GAACAGCTGGCGGACCGTCAAGGTGACCATCGGCCCAAGGCGCGCAAGCTGTAT 964  
 QY 131 TGAAGAACTTTTGTCTTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 190  
 DB 965 CGACAAGAAAGTTCGGCGCCCGACCATCAACCAACGAGCGGCTCACCATCGCCCGGT 1024  
 QY 191 CGAATTTAGAAGATCAATTTTGAAGCAATGGGAGCAAAATTTGTTGCTGAAGTGGCTTCAA 250  
 DB 1025 CGAGTGCAGCAGCGTACGAGAACCTCGGCGCCCGAGCTCGTCAAGAGGAGTGGCGACAA 1084  
 QY 251 AACCAATGATATGCTGGTATGGGACGACTACTGCAACAGTTTGTGACACAGGCAATGTT 310  
 DB 1085 GACCAACGACATCGGGGTGACGGCAACCAACCGGACCGGCTGTGCGCCAGGCGCTGTT 1144  
 QY 311 TCATGAAGGACTAAATAATGTCACAGAGGTGCTTAATCAATTTGTTATCCGTCGAGGAT 370  
 DB 1145 CGCGAGGCGCTGCGCAAGTTCGCGCGCGCGGCGCTCCCGCGCGCGCTGGAAGAGGGAT 1204  
 QY 371 TGAACAGCAACAGCAGCTGTTGAAGCCTTGAAGCCATTGTTCAACCTGTATCTG 430  
 DB 1205 CGAGCGCGCGCTGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCTGCGCGCGCTGCGCG 1264

QY 431 CAAGGAAGCTATTGCTCAGGTGCTGCTGAGTATCATCATCATCATCATCATCATCATCATCAT 490  
 DB 1265 CAAGTCCGACATCGCGCGCTGCGCGCGCTCTCCGCGCAGGACAAAGCAGTCTCGCGAGCT 1324  
 QY 491 TATCTCAGAAAGTATGGAGCGTGTGGGCAACGATGGTGTGATTTACCATCGAAGATCTCG 550  
 DB 1325 CATCGCCGAGCGATGGCAAGGTGCGAAGGACGAGTGTCTCAACGCTCGAGGATCCAA 1384  
 QY 551 AGGTATGGAACACAGAACTTGAAGTGTGTTGAAGCATGCAATTTGACCGTGGTACTCTGC 610  
 DB 1385 CACCTTCGGTGTGCGACTTCCAGCGGCGATGGGCTTCGACAGGGGTACTCTGTC 1444  
 QY 611 TCAATACATGTCACAGACAATGAAATAATGTTGACAGCTTGAATAATGTTTATCTT 670  
 DB 1445 CCGGTACATGTTGACGACGAGGATGAGGCGCTCTCGACGACCGCTACATCT 1504  
 QY 671 AATCAGGATATAAAGATGTCAAACATCCAGACATTTTGCACACTTGTGAGAGATTTCT 730  
 DB 1505 GATCCACGAGGCAAGATCGGTTTCGATCCAGACCTGTGTCGCGCTGCTGGAGAGTCA 1564  
 QY 731 TAAAC-----CAACCGTCCATTACTCATTTATTCAGATGATGTGATGTGAAGCACT 784  
 DB 1565 CGAGCGGGTGGCTCCAAGCCCTGCTGATCATGCCGAGGACGTCGAGGGCGAGGCCCT 1624  
 QY 785 TCCAAACCTTGTCTTGAACAAGATTCGTTGTTACTTTCAATGTTGTTGTTGTTGTTGTTGTT 844  
 DB 1625 GTCGACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1684  
 QY 845 AGGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 904  
 DB 1685 CGGCTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1744  
 QY 905 AGTGATTACAGAGGATAGGACTTGAATTAAGATGCTACAATGACAGCCCTTGGACA 964  
 DB 1745 GGTCTCTAAGATTACAGTTGATTAAGATGACAGCTAATTTGTTGAGGTTTCAAGAGTTTC 1024  
 QY 965 GGTCTCTAAGATTACAGTTGATTAAGATGACAGCTAATTTGTTGAGGTTTCAAGAGTTTC 1084  
 DB 1805 CGCGCGCGCGTCCCGTCCAGGAGGACGACGACCATCGTGGACCTGGAGAGGAGCGC 1864  
 QY 1025 AGAAGCTATTGCTAACCGTATTGCACTGATTAATTCGGAATTTAGAAACAAACAACTTCGA 1084  
 DB 1865 CGAGACCTTCCAGGCGCGCTGCGCCAGATCAAGCCGAGATTCGAGTCCGACGACTCGGA 1924  
 QY 1085 CTTTCAACGCTGAAATACAGAACTTGGCGAAATTTAGCTGTGTGTGTGTGTGTGTGTGT 1144  
 DB 1925 CTGGACCGCGAAGATTCAGGAGCGCTGCCAAGCTGGCCGCGCGCTGCTGCTGAT 1984  
 QY 1145 CAAAGTAGGAGCTCCACAGAGACAGCTTTTAAAGAAATGAAACTTCGATTGAGGATGC 1204  
 DB 1985 CGCGTCCGCGCGCGCGCGCGCTGAGGAGCGAGTGAAGGAGCGCAGCACCCTCTGGAGAGCGC 2044  
 QY 1205 TCTAATGCTACAGTGCAGCGCTTGAAGAGGATTCGTTGCTGTGTGTGTGTGTGTGTGTGT 1264  
 DB 2045 CATCTCCGCGCGCGCGCGCGCTGAGGAGGATTCGTTGCTGTGTGTGTGTGTGTGTGTGT 2104  
 QY 1265 TAT---TACGGTATTGAAAGTAGCAGCTTTCAGCTTTCAGGCGGATGATGCTACTGG 1321  
 DB 2105 GGTCCACCGCTCAAGTTCCTGGAGCAACCTCGGCGCGCAGCGCGAGAGCGCCACCGG 2164  
 QY 1322 ACGTAACATTTGCTTCTGTCTAGAGAGCGCTGTAGCTCAAAATGCTTTTAAATGCTGG 1381  
 DB 2165 TGTGCGGTGCTGCGCGCGCGCGCTGAGCGCTGCGCTGAGTTCGCGAGAGAGCGCGG 2224  
 QY 1382 GTACGAAGGCTCCGTAGTATTGACAGTTCAAAACAGCCCTGACGAGAACAGGATTTAA 1441  
 DB 2225 CTTGAGGCTACGTTCATCACCACCAAGTGGCGGAGCTGCACAAAGGCGCGGCTTCAA 2284  
 QY 1442 TGCTGCAACAGGTGAGTGGGTGATATGATTAACAGGAAATCATTTGACCTGTCAAGT 1501  
 DB 2285 CGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCGG 2344



Db 1201 TCCGCGACCCCGCGGTCGAGAGGGCATCGTCTCCGGTGGTGGCTCCGCGCT 1256

Search completed: February 15, 2001, 23:17:06  
Job time: 8581 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 22:00:13 ; Search time 925.81 Seconds  
(without alignments)  
12572.166 Million cell updates/sec

Title: US-09-001-737-7  
Perfect score: 1661  
Sequence: 1 GAATTCGGCTTCATATGGCA.....TGGGGGATAGCCGAATTC 1661

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues  
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*  
12: gb\_est12.\*  
13: gb\_est13.\*  
14: gb\_est14.\*  
15: gb\_est15.\*  
16: gb\_est16.\*  
17: gb\_est17.\*  
18: gb\_est18.\*  
19: gb\_est19.\*  
20: gb\_est20.\*  
21: gb\_est21.\*  
22: gb\_est22.\*  
23: gb\_est23.\*  
24: gb\_est24.\*  
25: gb\_est25.\*  
26: gb\_est26.\*  
27: gb\_est27.\*  
28: gb\_est28.\*  
29: gb\_est29.\*  
30: gb\_est30.\*  
31: gb\_est31.\*  
32: gb\_est32.\*  
33: gb\_est33.\*  
34: gb\_est34.\*  
35: gb\_est35.\*  
36: gb\_est36.\*  
37: gb\_est37.\*  
38: gb\_est38.\*  
39: gb\_est39.\*  
40: gb\_est40.\*  
41: em\_estba.\*  
42: em\_estfun.\*  
43: em\_esthum1.\*

44: em\_esthum2.\*  
45: em\_esthum3.\*  
46: em\_esthum4.\*  
47: em\_esthum5.\*  
48: em\_esthum6.\*  
49: em\_esthum7.\*  
50: em\_esthum8.\*  
51: em\_esthum9.\*  
52: em\_esthum10.\*  
53: em\_esthum11.\*  
54: em\_esthum12.\*  
55: em\_esthum13.\*  
56: em\_esthum14.\*  
57: em\_esthum15.\*  
58: em\_esthum16.\*  
59: em\_esthum17.\*  
60: em\_esthum18.\*  
61: em\_esthum19.\*  
62: em\_esthum20.\*  
63: em\_estin1.\*  
64: em\_estin2.\*  
65: em\_estin3.\*  
66: em\_estin4.\*  
67: em\_estov1.\*  
68: em\_estov2.\*  
69: em\_estpl1.\*  
70: em\_estpl2.\*  
71: em\_estpl3.\*  
72: em\_estpl4.\*  
73: em\_estpl5.\*  
74: em\_estrol.\*  
75: em\_estro2.\*  
76: em\_estro3.\*  
77: em\_estro4.\*  
78: em\_estro5.\*  
79: em\_estro6.\*  
80: em\_estro7.\*  
81: em\_estro8.\*  
82: em\_estro9.\*  
83: em\_estro10.\*  
84: em\_estro11.\*  
85: em\_estro12.\*  
86: em\_estro13.\*  
87: gb\_est41.\*  
88: gb\_est42.\*  
89: gb\_est43.\*  
90: gb\_est44.\*  
91: gb\_est45.\*  
92: gb\_est46.\*  
93: gb\_est47.\*  
94: gb\_est48.\*  
95: gb\_est49.\*  
96: gb\_est50.\*  
97: gb\_est51.\*  
98: gb\_est52.\*  
99: gb\_est53.\*  
100: gb\_est54.\*  
101: gb\_est55.\*  
102: gb\_est56.\*  
103: gb\_est57.\*  
104: gb\_est67.\*  
105: gb\_est68.\*  
106: gb\_est69.\*  
107: gb\_est70.\*  
108: gb\_est71.\*  
109: gb\_est72.\*  
110: gb\_est73.\*  
111: gb\_est74.\*  
112: em\_esthum21.\*  
113: em\_esthum22.\*  
114: em\_esthum23.\*  
115: em\_estom1.\*  
116: em\_estom2.\*

117: em\_estpl6:\*  
 118: em\_estpl7:\*  
 119: em\_estpl8:\*  
 120: em\_estpl9:\*  
 121: em\_estpl10:\*  
 122: em\_estpl11:\*  
 123: em\_estpl12:\*  
 124: em\_estpl13:\*  
 125: em\_estpl14:\*  
 126: em\_estpl15:\*  
 127: em\_estpl16:\*  
 128: em\_estpl17:\*  
 129: em\_estpl18:\*  
 130: em\_estpl19:\*  
 131: em\_estpl20:\*  
 132: em\_estpl21:\*  
 133: em\_estpl22:\*  
 134: em\_estpl23:\*  
 135: em\_estpl24:\*  
 136: em\_estpl25:\*  
 137: em\_estpl26:\*  
 138: em\_estpl27:\*  
 139: em\_estpl28:\*  
 140: em\_estpl29:\*  
 141: em\_estpl30:\*  
 142: em\_estpl31:\*  
 143: em\_estpl32:\*  
 144: em\_estpl33:\*  
 145: em\_estpl34:\*  
 146: em\_estpl35:\*  
 147: em\_estpl36:\*  
 148: em\_estpl37:\*  
 149: em\_estpl38:\*  
 150: em\_estpl39:\*  
 151: em\_estpl40:\*  
 152: em\_estpl41:\*  
 153: em\_estpl42:\*  
 154: em\_estpl43:\*  
 155: em\_estpl44:\*  
 156: em\_estpl45:\*  
 157: em\_estpl46:\*  
 158: em\_estpl47:\*  
 159: em\_estpl48:\*  
 160: em\_estpl49:\*  
 161: em\_estpl50:\*  
 162: em\_estpl51:\*  
 163: em\_estpl52:\*  
 164: em\_estpl53:\*  
 165: em\_estpl54:\*  
 166: em\_estpl55:\*  
 167: em\_estpl56:\*  
 168: em\_estpl57:\*  
 169: em\_estpl58:\*  
 170: em\_estpl59:\*  
 171: em\_estpl60:\*  
 172: em\_estpl61:\*  
 173: em\_estpl62:\*  
 174: em\_estpl63:\*  
 175: em\_estpl64:\*  
 176: em\_estpl65:\*  
 177: em\_estpl66:\*  
 178: em\_estpl67:\*  
 179: em\_estpl68:\*  
 180: em\_estpl69:\*  
 181: em\_estpl70:\*  
 182: em\_estpl71:\*  
 183: em\_estpl72:\*  
 184: em\_estpl73:\*  
 185: em\_estpl74:\*  
 186: em\_estpl75:\*  
 187: em\_estpl76:\*  
 188: em\_estpl77:\*  
 189: em\_estpl78:\*

190: gb\_gss25:\*  
 191: gb\_gss26:\*  
 192: gb\_gss27:\*  
 193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	243	14.6	645	87	AW224051	AW224051 EST300862
2	239.4	14.4	716	140	D46006	D46006 RICS10372A
3	237.6	14.3	637	87	AW223768	AW223768 EST300579
4	236.2	14.2	668	90	AW398404	AW398404 EST298251
5	227.4	13.7	680	89	AW330455	AW330455 PFOVAFGB3
6	226	13.6	1351	108	BE422296	BE422296 HWM022CF
7	225.2	13.6	709	21	A1486676	A1486676 EST244998
8	222.4	13.4	825	110	BE642158	BE642158 Cr12_5_E2
9	220.4	13.3	598	26	A1895994	A1895994 EST265437
10	214.4	12.9	825	97	AW927061	AW927061 HVSMEG000
11	211.2	12.7	695	106	BE318972	BE318972 NF043C12L
12	210.8	12.7	710	104	BE131653	BE131653 L48-1652T
13	209.4	12.6	612	87	AW216817	AW216817 EST295531
14	208.8	12.6	902	105	BE231179	BE231179 HVSMEG001
15	207.4	12.5	739	93	AW622566	AW622566 EST313366
16	201	12.1	677	108	BE435722	BE435722 EST406800
17	199.2	12.0	706	29	AU092993	AU092993 AU092993
18	197.8	11.9	578	108	BE432288	BE432288 EST398817
19	195.8	11.8	651	106	BE318418	BE318418 NF038802L
20	195.6	11.8	552	108	BE434127	BE434127 EST405205
21	195.2	11.8	730	137	BE901181	BE901181 60167579
22	195.2	11.8	819	110	BE643470	BE643470 Cr12_8_O1
23	194	11.7	621	97	AW934661	AW934661 EST353553
24	192.8	11.6	591	108	BE433229	BE433229 EST395758
25	191.8	11.5	557	93	AW618606	AW618606 EST320592
26	190.8	11.5	970	28	AU010446	AU010446 AU010446
27	190	11.4	546	39	AW030594	AW030594 EST273849
28	188.8	11.4	634	97	AW934424	AW934424 EST360267
29	186.8	11.2	734	28	AJ398447	AJ398447 AJ398447
30	184.4	11.1	491	108	BE435362	BE435362 EST406440
31	181.2	10.9	600	108	BE432172	BE432172 EST398701
32	181.2	10.9	886	135	BE799053	BE799053 601588241
33	180.8	10.9	584	38	AW004169	AW004169 701515092
34	180.2	10.8	494	108	BE471174	BE471174 WHE0285_G
35	180	10.8	840	135	BE779875	BE779875 601467908
36	180	10.8	1119	134	BE036873	BE036873 MP07G03_M
37	179.2	10.8	773	109	BE535246	BE535246 601059776
38	178.2	10.7	648	110	BE580002	BE580002 kg35f01.y
39	178	10.7	831	40	AW107008	AW107008 um18d11.y
40	177.2	10.7	609	137	BE918978	BE918978 FMI_2_E09
41	177.2	10.7	700	91	AW506070	AW506070 GEI379_G1
42	176.8	10.6	572	109	BE496298	BE496298 NXC1_022
43	176.4	10.6	578	108	BE432591	BE432591 EST399120
44	176	10.6	528	37	AV525715	AV525715 AV525715
45	176	10.6	828	106	BE300379	BE300379 600944444

# ALIGNMENTS

RESULT 1  
 AW224051 AW224051 645 bp mRNA EST 07-DEC-1999  
 LOCUS EST300862 tomato fruit red ripe, TAMU Lycopersicon esculentum CDNA  
 DEFINITION clone cLEN14H9, mRNA sequence.  
 ACCESSION AW224051  
 VERSION AW224051.1 GI:6535735  
 KEYWORDS EST.  
 SOURCE tomato.

ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids  
I; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 645)

REFERENCE Alcala J., Vrebalov, J., White, R., Matern, A.L., Holt, I. F., Liang, F.,  
Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,  
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)

AUTHORS Contact: David Frisch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dfrisch@CLEMSON.EDU

TITLE 5 prime sequence.  
JOURNAL Location/Qualifiers  
COMMENT 1. .645  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEN14H9"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/note=Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopene accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

205 a 102 C 166 g 172 t

BASE COUNT  
ORIGIN

Query Match	14.6%	Score 243;	DB 87;	Length 645;
Best Local Similarity	61.1%	Pred. No. 1.8e-57;		
Matches 393; Conservative	0;	Mismatches 250;	Indels	Gaps
QY	607	TGTCCTCAATACATGGTCACAGACAATGAAAAATGGTTGCAGACCTTGAAACCCCATTTA	666	
Db	2	TCCTCTCTACTTGTGTACCGACAGTCGGAATAATGTCGGTTGAATATGAGAACTGTTAAGT	61	
QY	667	TCCTTAATCAGGGATAAAAAAGTGTCAACATCCAAAGACATTTTCCCACTACTTCCAGGAAG	726	
Db	62	TGCTACTGGTTGTATAAAAAAGATACAAATGCAAGAGATCTGTGAATGCTCGTGAAGATG	121	
QY	727	TTCTTAAACCAACCGTCCATTTACTCATTTATTGTCAGATGTCGTGATGGTGAACCACTTC	786	
Db	122	CTATCAGAAATGGTTACCCAAATTTAAATATTGCTGAAGATATTGACGAGGAAGCTTTGG	181	
QY	787	CACCCCTGTCTTGAACAAGATTCGTGGTACTTTCAATGGTTGGTCTCAAAAGCGCCAG	846	
Db	182	CAACTCTTGTGTCAATAAAGCTTAGAGGTGCTTGAAGGTGCGTGCACATTAAAGCTCCTG	241	
QY	847	GATTTGGTGATCGTCGTAAGCTATGCTTGAAGACATTCGTATCTTCACAGGTGGTACAG	906	
Db	242	GTTTTGGTGAGCGNAAGAAGCCAGTATCTTGATGACATAGCAACCCCTTACTGGAGGCACTG	301	
QY	907	TGATTACAGAGGATCTAGGACTTGAATTTAAAGATGCTACAATGACAGCCCTTGGACAGG	966	
Db	302	TTATTAGGGAGGAGCTTGGGCTTACC'TTGGACAAGGCTGACAAGGAAGTTCTAGTCAATG	361	
QY	967	CTGTAGATATACAGTTGATAAAGNTAGCACAGTAATTTGTTGAAGGTTTCAGGAAGTTCAG	1026	
Db	362	CTGTAAAGTAGTGCTGACTAAGGATGCCACTACAATTTGTGGTGTGGTAGCACTCAGG	421	
QY	1027	AAGCTATTGCTAACCGTATTTCACATTAATCCCAATTTAGAAACAACAACTTCTCACT	1086	

Db	422	AAGCAGTCAACAAACGTTGTTGCACAGATTAAAAACCTCATAGAGCGTGCAGATCAAGATT	481
Qy	1087	TTGACCGTGAAAAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCA	1146
Db	482	ATGAAGAGAAAAGCTAAATGAAGAAGATTGCTAAAGTTATCAGGAGGTGGCTGTCATAC	541
Qy	1147	AAGTAGGAGCTCCAACAGACAGACGCTTTAAAGAAATGAAACTTCGCATTGAGGATGCTC	1206
Db	542	AGTTGGAGCTCAAACTGAACTGAATTTGAAGGAGAGAAGAACTTAGAGTAGAAGATGCTC	601
Qy	1207	TAAATGCTACAGCTGCAGCCGTTGGAAGAAGTATCGTTGCTGG	1249
Db	602	TCAATGCAACAAGACGAGCTGTTGAGGAAGGATTTGTTGTTGG	644
RESULT 2			
D46006	D46006	716 bp	EST
LOCUS	RICS10372A	Rice green shoot	Oryza sativa cDNA, mRNA sequence.
DEFINITION	D46006		
ACCESSION	D46006.2	GI:7212768	
VERSION			
KEYWORDS	EST.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.		
AUTHORS	1 (bases 1 to 716)		
TITLE	Sasaki, I., Miyao, A. and Yamamoto, K.		
JOURNAL	Rice cDNA from callus 1995		
COMMENT	Unpublished (1995)		
	On Mar 9, 1995 this sequence version replaced gi:699715.		
	Contact: Takuji Sasaki		
	National Institute of Agrobiological Resources		
	Rice Genome Research Program		
	1-1-2 Kannondai, Tsukuba		
	Ibaraki,		
	Japan 305		
	Tel: 0298-38-7441		
	Fax: 0298-38-7468		
	Email: tsasaki@agr.affrc.go.jp		
	PROJECT = 'RGP'		
FEATURES	Sequence updated (01-Mar-2000).		
source	Location/Qualifiers		
	1..716		
	/organism="Oryza sativa"		
	/strain="Nipponbare"		
	/db_xref="taxon:4530"		
	/clone_lib="Rice green shoot"		
	/note="Green shoot (8 days old)"		
BASE COUNT	222 a 120 c 198 g 175 t		1 others
ORIGIN			
Query Match	14.4%;	Score 239.4;	DB 140; Length 716;
Best Local Similarity	58.6%;	pred. No. 1.9e-56;	
Matches	414; Conservative	0; Mismatches 292;	Indels 0; Gaps 0;
Qy	550	GAGGTATGGAACAGAACTTGAAGTGGTTGAGCGATGCAATTGACCGTGGTTACCTGT	609
Db	5	CAAGTCTTCAGAAATAATCTCTATCTTGTGGAGGGAATGCGAGTTTGAACGTTGGTTATATCT	64
Qy	610	CTCAATCATNGTCCACAGACAAATGAAAATGTTGTCAGACCTTGAACCCATTTATCT	669
Db	65	CCCGTATTTTGTAACTGACAGCGAGAAATGCTCGAGAAATATGAGAACTGCAANCTTC	124
Qy	670	TAAATCAGGATATAAAAGTGTCAAACTCCAAGACATTTTGCCACTACTTTGAGGAAGTTC	729
Db	125	TTTTGGTGACAAAAAATCACCACCGAGGATCTTATCAATGTTTGGAGAGGCCA	184
Qy	730	TAAACCAACCGTCCATTACTTATTTCAGATGATGGATGGTGAAGCACTTCCAA	789
Db	185	TAAAGGTGTCATACCCAAATCTTGATATGCTGAGGATATTGACGAGGAGGCTTGTGCTA	244





QY 611 TCAATACATGGTCACAGACAATGAAAAAATGTTGCGAGACCTTTGAAAAACCCATTTATCTT 670  
 Db 1 TCGTATTTTATTACAAATAATGAGAAATGATTGGAGCTTGATGATCCATATCTCT 60  
 QY 671 AATCAGGATAAAAAGTGTCAACATCAACAGACATTTTCCACACTTCTTGAGGAAGTTCT 730  
 Db 61 AATACAGAGAAAAAATCTTAATATTATTCAGGCTTTGCTTTCTTATTTCTTGAAGCTGTCT 120  
 QY 731 TAAACCAACCGTCCATTTACTTATTCAGATGATGTTGAGTGTGAAGCACTTCCCAAC 790  
 Db 121 TAAGTCTGTAACCTTTACTTATTCATTCGCAAGACATTTGAGGTTGAAGCATTAAGTAC 180  
 QY 791 CTTGCTTTGAACAAGATTCGTGGTACTTTCAATGTTGTTGCTGTCACAGCCGAGGATT 850  
 Db 181 TTTAGTTATTATAAGTTCGTTGGAGTCTAAAAGTTGCTGCAGTGAAGCTCCAGGTTT 240  
 QY 851 TGGTATCGTCTGAAGCTATGCTTTGAAGACATTTGCTATCTTTGACAGGTTGTA---CAGT 907  
 Db 241 TGGTGATAGAAGAAAGGAGATGCTTTGAAGATATAGCAGCTTTAACTTAATGCTAAAGTATCT 300  
 QY 908 GATTACAGAGGATCTAGGACTTTGAATTTAAAGATGCTACAATGACAGCCCTTTGGACAGGC 967  
 Db 301 CATAAAGATGAACCTTTGGATTAAATGCAAGACCTTAACCTTTGAAGACCTTTGGCATTCG 360  
 QY 968 TGTAAAGATTACAGTTGATAAGATAGACAGATGATTTGTTGAAGTTTCAGGAAGTTTCAGA 1027  
 Db 361 TAAAAATGTTAAATCACTAAAAGATAATACTACAATTTGTTAGGAAATAGAGTTACTGA 420  
 QY 1028 AGCTATTGCTAACCTTTATGCACTGATTAATTCGCAATTTAGAACACAACTTTCGACTT 1087  
 Db 421 CAGAGTAAAGAGCTAGATTTGACAGATTAATTTCAATTTGAGTCTTCAACTTTCTGATTA 480  
 QY 1088 TGACCGTGAAGAACTACAGAAAGCTTTGGCGAAATTAGCTGCTGGTGTAGCTTTATCAAA 1147  
 Db 481 TGTAAAGAGAATTAAGAGAGCTTTTACCAAAATTAATCAGTGGTGTGCTTACTATAA 540  
 QY 1148 AGTAGAGCTTCCAAGAGACAGCTTTTAAAGAAATGAAACTTCCGATTTAGAGTGTCTT 1207  
 Db 541 AGTTGTTGAGCAGCACTGAAATAGAGTTTAAAGAACCTTAAAGATAGGTCGAGGATGCTT 600  
 QY 1208 AATGCTACAGCTGACCGCTTTGAAGAAG 1236  
 Db 601 GCACGCCACAAGAGCTGCAATTTGAGGAAG 629

## RESULT 6

LOCUS BE422296 1351 bp mRNA EST 24-JUL-2000  
 DEFINITION HW022cf.04r ITEC HWM Barley Leaf Library Hordeum vulgare cDNA  
 clone HW022cf.04, mRNA sequence.  
 ACCESSION BE422296  
 VERSION BE422296.1 GI:9420139  
 KEYWORDS EST.  
 SOURCE Hordeum vulgare  
 ORGANISM  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.  
 REFERENCE 1 (bases 1 to 1351)  
 AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,  
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,  
 Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,  
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,  
 Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M.,  
 Sorrells, M., Warburton, M. and Wenzel, G.  
 TITLE International Triticeae EST Cooperative (ITEC): Production of  
 Expressed Sequence Tags for Species of the Triticeae  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Wenzel G  
 TU Muenchen, Lehrstuhl fuer Pflanzenbau und Pflanzenzuchtung  
 Am Hochanger 2, D-85350 Freising-Weihenstephan GERMANY  
 Fax: 49 08161 71 5173  
 Email: wenzel@em.pbz.agrar.tu-muenchen.de

International Triticeae EST Cooperative (ITEC)  
 http://wheat.pw.usda.gov/genome.

FEATURES  
source

1. 1351  
 /organism="Hordeum vulgare"  
 /cultivar="Barke"  
 /db\_xref="taxon:4513"  
 /clone="HW022cf.04"  
 /clone\_lib="ITEC HWM Barley Leaf Library"  
 /tissue\_type="leaf"  
 /dev\_stage="14 day old"  
 /note="Vector: pBluescriptSK(-); 850 bp average insert  
 size."

BASE COUNT 376 a 281 c 372 g 317 t 5 others  
 ORIGIN

Query Match 13.6%; Score 226; DB 108; Length 1351;  
 Best Local Similarity 58.6%; Pred. No. 1.5e-52;  
 Matches 391; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 632 TCAAAAAATGGTTGCAGACCTTTGAAACCCATTTATCTTAATCACGGATAAAAAGTGTG 691  
 Db 119 TGAGAAATGACCCAGGTACGAGAACTGCAAGCTGCCTTTGGTTGCAAGAAATAC 178  
 QY 692 AAACATCCAAAGACATTTTGCCTACTTCTGAGGAAGTTCTTAAACCAACCGTCCCTACT 751  
 Db 179 CAACGACGGATCTTATCAATGTTCTGGAGGACCATTAGGGTCAATACCACTCT 238  
 QY 752 CATATTGTCAGATGATGGATGGTGAAGCACTTCCAAACCTTGTCTTGAACAAGATTCG 811  
 Db 239 GATCATTTGCTAGGATATTGAGCAGGAGCTCTTGCACCCCTTGTGTCAACAAGCTAAG 298  
 QY 812 TGSTACTTTCAATGGTGTGCTGTCNAACGCCAGGATTTGGTGTGCTGCTAAGCTAT 871  
 Db 299 AGTTCTTTGAAATCTGTGCTATCAAGCCCTTGGTTTGGAGAGCGCAACCCAGTA 358  
 QY 872 GCTTGAAGACATTTGCTTGTACAGGTGTTACAGTGATTACAGAGGATCTAGGACTTGA 931  
 Db 359 CTTGACGACATTTGCCATCTCACCAGGAGAACTGTTATCAGAGAGAGGTTGGACTCAC 418  
 QY 932 ATTAAGATGTCATACATGACAGCCCTTGGCAGGCTGCTAAGATTACAGTTGATAAGA 991  
 Db 419 ACTTGACAGGAGATTAACACAGTTCTAGGAACGGCTGCAAGAGTTGTCTTACAAAAGA 478  
 QY 992 TAGCACAGTAATTTGTTGAAGTTTCAGGAAGTTTCAAGAGCTATTGCTTAACCGTATGCACT 1051  
 Db 479 GTCGACACAATAGTTGGTGTGCGCAGCACCCAGGAAGAAGTGAAGAGGTTGCACA 538  
 QY 1052 GATTAATCGCAATTTAGAAACAACTTCTGACTTTTGACCGTGAAAACTACAAGAAG 1111  
 Db 539 GATCAAAAATCTCATTTAGGTAGCAGAGCAACACTACGAGAGGAAAGAACTCAATGAGAG 598  
 QY 1112 TTTGGCGAAATTTAGCTGGTGGTGTAGCTGTTATCAAGTAGGAGCTTCCAAAGAGACAG 1171  
 Db 599 GATTCAAGAGCTCGCGGTTGTTGCTGTTATTAGTGGGAGGACCAACAAGAACTGA 658  
 QY 1172 TTTAAAGAAATGAACCTTCGCATTGAGATGCTCTAAATGCTACAGCTGACGCGTGA 1231  
 Db 659 ACTTAAGAGAGAAGATTGCGAGTTGAGATGCTCTANACGCAACCAAGGCTGCGCTGA 718  
 QY 1232 AGAAGTATCGTTGCTGGTGGTGGAAACAGCACTATTACGGTTATTGAAAAAGTAGCAGC 1291  
 Db 719 GGAAGTATTGTTGTTGGTGGAGGCTGACACTCTTTTGGAGCTGGCTGCTAAGTTGATGC 778  
 QY 1292 TCTTGAAG 1298  
 Db 779 CATCAAG 785

## RESULT 7

LOCUS AT486676 709 bp mRNA EST 29-JUN-1999

Db	408	AAGTCTCTGTTTTGGTGAGCGAAAAACCCAGTATCTTGTATGACATAGCAACCCCTTACTG	467
QY	898	GTGCTACAGTGATTACAGAGAGATCTAGGACTTGAATTTAAAGATGCTACAAATGACAGCCC	957
Db	468	GAGGCACGTGTATTAGGAGGAGGCTTGGCCCTTACCTTGGACAGGCTGACAAGGAAGTTC	527
QY	958	TTGGACAGGCTGCTAAGATTACAGTTGATATAAGATACACAGTAAATTTGTTGAAGGTTTCA	1017
Db	528	TAGGTCATGCTGCTAAAGTAGTGTCTAGTAAAGTATGAGGATGCCACTACATAATTTGTTGGTATGGTA	587
QY	1018	GAAGTTTCAGAGCTATTGCTTAACCGTATTGCACGTATTAAATCCCAATTAGAACACACAA	1077
Db	588	GCACTCAGGAAGCAGTCAACAAACGTTTGCACAGATTAACAAAGCTGATAGAGCTGCAG	647
QY	1078	CTTCTGACTTTTGGCCGTGAAAAACATACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAG	1137
Db	648	ATCAAGATTATGAAGAGAAAGCTAAATGAAGAAATGCTAAATTTATCAGGAAGTGTGG	707
QY	1138	CT 1139	
Db	708	CT 709	
RESULT 8			
BE642158			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
QY			
Db			
Query Match			
Best Local Similarity			
Matches			
QY			
Db			





KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

EST.  
barley.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.  
1 (bases 1 to 825)  
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo,T., Sasaki,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and Wood,T.  
Development of a genetically and physically anchored EST resource for barley genomics  
Unpublished (2000)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
High quality sequence stop: 825.  
Location/Qualifiers  
1..825  
/organism="Hordeum vulgare"  
/cultivar="Morex"  
/db\_xref="taxon:4513"  
/clone="HVSMEq0009G12"  
/clone\_lib="Hordeum vulgare pre-anthesis spike EST library HVCdNA0008 (white to yellow anther)"  
/tissue\_type="pre-anthesis spike"  
/lab\_host="SOLR"  
/notes="Vector: lambdaDazAP; Site\_1: EcoRI; Site\_2: XhoI"  
242 a 166 c 237 g 179 t 1 others

Query Match 12.9%; Score 214.4; DB 97; Length 825;  
Best Local Similarity 60.4%; Pred. No. 2.2e-49;  
Matches 353; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 715 TACTTTGAGGAAGTCTTTAAACCAACCGTCATTACTCATTTATTCAGAGCATGCTGGATG 774  
Db 11 TTTCTGGAGGAAGCATTAGGGGTCAATACCCAATCCCTGATCATTCGTGAGGATATTGAGC 70

QY 775 GTGAAGCACATCTCCAACCCCTTCTTTGAACAAGATTCTGTTACTTTCATTTCAATGTGGTCTG 834  
Db 71 AGGAGGCTCTTGCACACCCCTGTTGTCACACAGCTAAGAGGTTCTTTGAAATCTGTGCTA 130

QY 835 TCAAAGCCGACGAGATTGGTGATCGTCGTAAGCTATGCTTGAAGACATTGCTATCTTGA 894  
Db 131 TCAAAGCCCTCGTGTGTTGGAGAGGCCAAGCCACAGTACCTGGACGACATTGCCATCTCA 190

QY 895 CAGTGTGTACAGTGATTACAGAGGATCTAGGACTTGAATTAAGAGTGTCTACATGACAG 954  
Db 191 CCGGAGGAACCTGTATTACAGACAGAGGTTGGACTCACACTTGACAAAGCAGATAACACAG 250

QY 955 CCCTTGGACAGGCTGCCTAAAGATTACGTTGATAAAGATAGCACAGTAATTCCTTCAAGGTT 1014  
Db 251 TTCTAGGAACGGTGCACAAAGTTGTCCTTACAAAGAGTCGACACAACTAGTTCGTTGATG 310

QY 1015 CAGGAAGTTCAGAAGCTATTGCTAACCGTATTGCACCTCATTAATTCGCAATTTAGAAACAA 1074  
Db 311 GCAGCACCCAGGAAGAAGTGACTAAGAGGGTTGCACAGATCAAAAATCTCATTGAGGTAG 370

QY 1075 CAACTTCTGACTTTCACCGTGAACAACTACAGAACGTTTGGCGAATATTAGCTGGTGGTG 1134  
Db 371 CAGACGAAGACTACGAGAAGGAAAAACTCAATGAGAGATTGCAAAAGCTGCCCGTGGTG 430

QY 1135 TAGCTGTATTCAAGTAGTAGGAGCTCCAACAGACAGAGCTTTAAAGAAATGAAACTTCGCA 1194  
Db 431 TTGCTGTATTTCAGTGGGAGACACACAGAACTGAACTTAAGGAGAAGAGTTGGCAG 490

QY 1195 TTGAGTAGTCTTAATTCGTACAGTGCAGCCGTTGAAGAGGTATTCGTTGCTGGTGGTG 1254

```
Db 268 ATCTTGACAGGAGGTACTGTATTACAGAGAGGTTGGCTTAGCTTAGACAAAGCTGGG 327
QY 948 ATGACAGCCCTTGCACAGGCTGTAAGATTACAGTTGATAAAGATACACAGTAATTTGTT 1007
Db 328 ATGAGGTTCTAGGCACTGCGGCCAGGTTGGTCTCACCAGAGATACACACCAATAGTA 387
QY 1008 GAAGGTTTCAGGAAGTTGAGAAGCTATTGCTAACCGTATTGCACCTGATTAAATCGCAATTA 1067
Db 388 GGTGATGAAGTACCCAGAGAGCAGTTACCAACGGAGTTTCACAAATTAAGAACCRAATT 447
QY 1068 GAACAAACAACTTCTGACTTTGACCGTGAAAACTACAAGAGCTTTGCGGAAATTAGCT 1127
Db 448 GAGGCTGCAGAAAGACTATGAAAAGGAGAGCTGAACGAAAGGATTGCAAAACTGTCT 507
QY 1128 GGTGGTGTAGCTGTTATCAAGATGAGAGCTCCCAACAGACAGCTTTAAAGAAATGAA 1187
Db 508 GGTGGTGTGCTGTAATTCAGGTTGTGTGACAAACTGAGACAGATCTCAAGGAAAGATA 567
QY 1188 CTTCCGATTGAGGATGCTCTAAATGCTACACGTGCAGCGCTTGAAGAAGGTATCGTTGCT 1247
Db 568 TTGAGAGTTGAGGACGCTCTATATGCCACAAAGCGAGCTGTTGAAAGGATTTGTAGTT 627
QY 1248 GGTGGTGG 1255
Db 628 GCGCGGGG 635

RESULT 12
BE131653
LOCUS
DEFINITION
L48-1652T3 Ice plant Lambda Uni-zap XR expression library, 48 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-1652,
mRNA sequence.
ACCESSION
BE131653
VERSION
BE131653.1 GI:8579016
KEYWORDS
EST.
SOURCE
Mesembryanthemum crystallinum
common ice plant.
ORGANISM
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997).
JOURNAL
Contact: Cushman JC
Department of Biochemistry and Molecular Biology
Oklahoma State University
350 Noble Research Center, Stillwater, OK 74078-3035, USA
Tel: 405-744-6207
Fax: 405-744-7799
Email: jcushman@biochem.okstate.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-17 row: E column: 4
Seq primer: T3
High quality sequence stop: 350
POLYA-No. Location/Qualifiers
1. .710
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L48-1652"
/clone_lib="Ice plant Lambda Uni-zap XR expression library
, 48 hours NaCl treatment"
/tissue_type="Leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 225 a 122 c 194 g 169 t
```

```
ORIGIN
Query Match 12.7%; Score 210.8; DB 104; Length 710;
Best Local Similarity 58.0%; Pred. No. 2.2e-48;
Matches 391; Conservative 0; Mismatches 282; Indels 1; Gaps 1;
QY 582 GGCATGCAATTTGACCGTGGTTACCTGTCTCAATACATATGTCACAGCAATGAAAAATG 641
Db 1 GGAATGCAATTTGATCGAGCTATATCTCCCATATCTTTGTACACATAGTAGTAATAATG 60
QY 642 GTTGACAGCTTTGAAAACCCATTTATCTTAATCACGGTAAAAAGTGTCAAAACATCCAA 701
Db 61 ATTGTTGAATATCAAAACTGTAAGTTGCTTGGTAGACAGAAAGATCACAACTGCGAGA 120
QY 702 GACATTTTCCCACTACTTCAGGAGTTCCTTAACACCAACCGTCCATTAATCTACTTATTCGA 761
Db 121 GACCTAGTTAATGTCTAGAAAGATGCAATTAACCAACTACCCCAATCTTGATTAATCGCA 180
QY 762 GATGATGTGGATGGTGAAGCAGCTTCCAAACCCCTGTCTTGAACAAGATTCGTGTACTTTC 821
Db 181 GAAGACATTTGACGAGCAACCTCTTGCTACATTAAGTTGTAACAAGCTCCGGGATCACATT 240
QY 822 AATGTGTTGCTGTCAAAAGCGCAGGATTGTTGATGTCGTAAAGCTATGCTTGAAGAC 881
Db 241 AGGATAGCTGCAATGAAGGCTCGGGTTATGGAGATCGAAGCATCAATATCTTGTATGAT 300
QY 882 ATTGCTATCTTGACAGTGTGACAGTGAATACAGAGATCTAGGACTTGAATTAAGAAT 941
Db 301 ATTGCTATCTTACTGGAGGAACCTGTGATCAGAGAGAGGTTGTTGTTTGGACAAA 360
QY 942 GGTACAATGACAGCCCTTGGACAGGCTGTCTAGATTAACAGTGTGTAAGAGTAGCACAGTA 1001
Db 361 GTTGGCAGAGAGTCTTGGGACATGCTTCAAGGTTGCTCAAAAAGAGTCAGCTACC 420
QY 1002 ATTGTTGAAGTTTCAGGAAGTTTCAGAAAGTATTTGCTTAACCGTATTGCACCTGATTAATCG 1061
Db 421 ATTGTTGGGGTGGGAGCACACAAGAAGCAGTTACTAGGAAAATTTGCCAGATTAGGAAG 480
QY 1062 CAATTAGAAAACAACACTTCTGACITTGACCGTGAAGAACTACAAAGCTTTGGCGAAA 1121
Db 481 CAGCTGAGAGTGTGCTGACCCAGGATGACGAAAACAAGAGCTCAATGAAGGATAGGAAAA 540
QY 1122 TTAGCTGGTGGTGTAGCTGTATTCAAAAGTAGGAGTCCCAACAGAGACAGCTTTAAAGAA 1181
Db 541 CTCTCCAGCGGTGTGCTGTGATTTCAGGTTGGAGCCAGACTGAGACTGAGATGAGAGAA 600
QY 1182 ATGAACCTTCGCATTTGAGGATGCTCTTAATGCTACAGTGTGACGCCGTTGAAGAAGTATC 1241
Db 601 AAGAACTGAGAGTGGAGGATGCTCTTAATGC-ACAAAGGCAGCTGTAGATGAAGGTATC 659
QY 1242 GTTGGTGGTGGTGG 1255
Db 660 GGTGATGGTGGAGG 673

RESULT 13
AW216817
LOCUS
DEFINITION
EST295531 tomato callus, TMMU Lycopersicon esculentum cDNA clone
cLEC87D23, mRNA sequence.
ACCESSION
AW216817
VERSION
AW216817.1 GI:6527691
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 612)
AUTHORS
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Alm,S., Ronning
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
```

TITLE Generation of ESTs from tomato callus tissue  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: David Frisch  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 4366  
 Fax: 864 656 4293  
 Email: dfrisch@CLEMSON.EDU  
 5 prime sequence.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

1. .612
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC87D23"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLL-Blue MRF"
/note="Vector: pBluescript SK(-); Source:
xhol; supplier: Giovannoni laboratories
of seedlings 7-10 days post-germination
at both ends and placed on MS medium.
Mixed callus was harvested at 25 and
undifferentiated masses. Tomato callus
190 a 95 c 160 g 166 t 1

```

BASE COUNT  
ORIGIN

Query Match 12.6%; Score 209.4; DB 87; Length 612;  
Best Local Similarity 58.8%; Pred. No. 5.1e-48;  
Matches 360; Conservative 0; Mismatches 252; Indels 0;

QY 513 GTGGCAACGATGGTGTGATTACCATCGAAGAAATCTCGAGGTATGGAAACAGAACTTGAA 572

Db 1 GTTGAAGGAAGTGTGTGACACTAGAGGAGGGTAAAAGTCTGAAAAACAGTCTGCGT 60

QY 573 GTGGTTGAAGGCATGCAATTGACCGTGGTTACCTGTCTCAATACATGGTCACAGACAAT 632

Db 61 GTGGTTGAAGGAATGCAATTGACCGTGGTTATGTCTCTCCTTACTTTGTACCGACAGT 120

QY 633 GAAAAATGGTTCAGACCTTGAAAACCCATTATCTTAATCACGGATAAAAAAGTGTC 692

Db 121 GAGAAAATGTCCGTTGAATATGAGAACTGTAAAGTTGCTACTGGTTGATAAAAAGATAACA 180

QY 693 AACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCCTTAAACCAACCGTCCATTACTC 752

Db 181 AATGCAAGAGATCTTGTAAATGTCTCTGGAAGATGCTATCAGAAATGGTTACCCAATTTA 240

QY 753 ATTATTCAGATGATGGATGCTGAAGCACCTTCCAACCCCTTGTCCTTGAACAAGATTTCGT 812

Db 241 ATATTGCTGAAGATATTGAGCAGGAAGCTTTGGCAACTCTTGTTGTCAATAAGCTTAGA 300

QY 813 GGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGATTTGGTGATCGTCGTAAGCTATG 872

Db 301 GGTGCCCTTGAAAGGTCGCTGCACCTTAAAGCTCCTGGTTTGGTGAGCGGAAAAGCCAGTAT 360

QY 873 CTTGAAGACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAGGATCTAGGACTTGAA 932

Db 361 CTTGATGACATAGCAACCCCTTACTGGAGGCACCTGTTATTAGGGAGGAGCTTGGCCTTACC 420

QY 933 TTAAAGATGCTACAATGACAGCCCTTGGACAGGCTGAAGATTACAGTTGATAAAGAT 992

Db 421 TTGGACAAGGCTGACAAGGAAGTTCTAGGTCATGCTGCTAAAGTAGTGCTGACTAAGGAT 480

QY 993 AGCACAGTAATTGTTGAAGGTTCAGGAAGTTCAGAAGCTATTGCTAACCGTATTGCACTG 105

Db 481 GCCACTACAATTGTTGGTGATGGTAGCACTCAGGAAGCAGTCAACANACGTGTTGCACAG 540

QY 1053 ATTAATCGCAATTAGAAACAACAACCTCTGACTTTGACCGTGAAAACTACAAGAACGT 111

Db 54.1 ATTAACCTGATAGAGCTGCAGATCAAGATTATGAAAAGGAAAAGCTAAATGAAAGA 600

Qy	1113	TTGGCGAAATTA	1124
Db	601	ATTGCTAAGTTA	612

RESULT 14

BE231179	BE231179	902 bp	mrna	EST	25-JUL-2000
LOCUS	HY5MEG0017G14f	Hordeum vulgare	pre-anthesis	spike EST library	
DEFINITION	HYCDNA0008 (white to yellow anther)	Hordeum vulgare	cDNA clone		
	HY5MEG0017G14f, mRNA sequence.				
ACCESSION	BE231179				
VERSION	BE231179.1	GI:8967402			
KEYWORDS	EST.				
SOURCE	barley.				
ORGANISM	Hordeum vulgare				

## REFERENCE

WING R., Close, T. J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu  
Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo  
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D. W., Main, D. and  
Wood, T.

**TITLE**

JOURNAL  
COMMENT  
for barley genomics  
Unpublished (2000)  
Contact: Wing RA  
E-mail: wing@alum.mit.edu

## COMENI

Clemson University Economics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293

```

Seq primer: AATTAAACCCCTCACTAAAGGG
High quality sequence start: 7
High quality sequence stop: 578.
Location/Qualifiers
    1..902
        /organism="Hordeum vulgare"
        /cultivar="Morex"
        /db_xref="taxon:4513"
        /clone="HVSMEg0017G14f"
        /clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCNDA0008 (white to yellow anther)"
        /tissue_type="pre-anthesis spike"
        /lab_host="SOLR"
        /note="vector: lambda2AP; Site_1: EcoRI; Site_2: XhoI"
256 a 199 c 250 q 196 t 1 others
BASE COUNT

```

BASE COUNT  
ORIGIN

Query Match	12.6%	Score 208.8;	DB 105;	Length 902;
Best Local Similarity	57.6%	Pred. No. 8.8e-48;		
Matches 392:	Conservative	0: Mismatches 288:	Indels :	Gaps 1:

QY 587 GCAATTGACCGTGGTTACCTGTCTCAATACATGTGTCACAGACAATGAAAAATGGTTGC 646

Db 9 GGAGATTGACCGTGGCTACATTCCCCCTCAATTTCGTGACAAACCTTGAGAAATCGATTGT 68

QY 647 AGACCTTGAAAACCCCATTTATCTTAATCACGGATAAAAAAGTGTCAAAACATCCAAGACAT 706

Db 69 GGAGTTTGAGAAATGCTAGAGTTCTTATAACTGATCAGAAGATCACAGCATAAAGGAAAT 128

QY 707 TTTGCCACTACTTGAGGAAGTTCTTAAACCAACCGTCCATTCA'TTATTGCAGATGA 766

Db 129 CATTCCACTTCTGGAGCAGACTACACAGTTGAGATGTCCGCTATTCA TTGTAGCCGAGGA 188

QY 767 TGTGGATGGTGAAGCACTTCCAACCCCTTGTC TTGAACAAGATTCTGGTGGTACTTTCAATGT 826

Db 189 CATTACTGGTGAAGCTTTGGCGACTCTTGTGTAAACAAGCTCAGAGGTATTATTACGT 248

QY 827 GGTTCGCTCAAAGCGCCAGGATTTGGTGATCGTCGTAAGCTATGCTTGAAGACATTGC 886

```

Db 249 TCGCGCATCAAAAGCCCAAGTTTCGGTGAGCGCGGAAGGCTGTCTTCAGGATATTC 308
QY 887 TATCTTGACAGTGGTACAGTATTACAGAGGATCTAGGACTTGAATTTAAAGATCTAC 946
Db 309 CATCGTGACAGTGGTGAATACCTACGGAAGGATCTTGGTCTGGTTGAGAAATCGGAC 368
QY 947 AATGACAGCCCTTGGACAGGCTGCTAAGATTACAGTTGATAAGATGACAGTAAATGCT 1006
Db 369 GGTAGACCAACTTGGGACAGCAAGGAAATCAATTTATCATGACACTACAACACCCCTCAT 428
QY 1007 TGAAGCTTTCAGGAAGTTCAGAAGCTATTGCTTAACCGTATTGCACTGATTTAAATCGCAATT 1066
Db 429 AGCAGATGACAGTACGAGCAAGAGAGATCCAGCGAGGGTTGACAGCTAAAGAAAGAGCT 488
QY 1067 AGAAACAACAACCTTTCGACTTTGACCGTGAANAACCTACAAACGCTTTGGCGAAATTTAGC 1126
Db 489 TTCTGAACTGATTCATCTATGATCTCAGAAATTTGGCTGAGAGATTTGCCAAGCTTTC 548
QY 1127 TGGTGTGTAGCTGTTTATCAAGTAGGAGCTCCACAGACAGAGCTTTTAAAGAAATGAA 1186
Db 549 TGGTGTGTGGCGTCAATCAAGGTTGGAGNACAACTGAGACAGAGCTCCGAGACCGTCA 608
QY 1187 ACTTCCATTGAGGATGCTCTAAATCTACAGCTGCAGCGCTTGAAGAAGGTATCGTTCC 1246
Db 609 GCTGCGGATCGAGGAGCGGAAGACCCACTTTCGCTG-CATCGAGGAAGGCAATTGGTCC 667
QY 1247 TGGTGTGGACAGACACTTAT 1267
Db 668 CGGTGTGTGGCGGATATGT 688

RESULT 15
AW622566 739 bp mRNA EST 28-MAR-2000
LOCUS ESR313366 tomato root during/after fruit set, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cLEX15J21 5', mRNA sequence.
ACCESSION AW622566
VERSION AW622566.1 GI:7334213
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 739)
AUTHORS van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Kochian,L.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato root, during and after fruit set
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
1..739
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX15J21"
/clone_lib="tomato root during/after fruit set, Cornell
University"
/tissue_type="root"
/dev_stages="plants during and after fruit-set"
/notes="Vector: pBluescript SK(-); Site:1: EcoRI; Site:2:
XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."
```

Search completed: February 16, 2001, 01:02:10  
Job time: 10917 sec

```

BASE COUNT 234 a 125 c 200 g 179 t 1 others
ORIGIN

Query Match 12.5%; Score 207.4; DB 93; Length 739;
Best Local Similarity 58.2%; Pred. No. 2e-47;
Matches 383; Conservative 0; Mismatches 272; Indels 3; Gaps 1;

QY 13 ATATGGCAAAACAAATTTTCAGCAGATGCGGCTGCTGCCATGTTGCGCGAGTTG 72
Db 78 ATGCTGCTAAGGATATAATTTGGAGTTGAAGTAGGGGAATTATGCTTCAAGAGTTG 137
QY 73 ATATGTTAGCAGATACCGTCAAGTAACGCTTGGTCTTAAAGGGGCAATGTTGTTG 132
Db 138 AGCAGCTTCTGATGCAAGTCAAAAGTCACTATGGTCCAAAGGTCGTAATGTTGTTG 197
QY 133 AAAAGCTTTTGGTCTCCCTTAATTAATTAATGACGGGTAACCATGCTTAAGAGATCG 192
Db 198 ACAAGTTTGGGGTGCACCCCAAGGTGACAAAAGATGGTGTCACTGTTGCCAAAAGCAT 257
QY 193 AATTAGAAGATCATTTTGAACACATGGGAGCAAAATTTGGTCTCTGAAGTGGCTTCTAAA 252
Db 258 AATTCAAGGATAAATAAATAATTTGGTCCAGCCTTTGAAGCAAGTTGCCAATGCCA 317
QY 253 CCAATGATATTGCTGGTATGGGACGACTACTGCAACAGTTTGGACACAAGCCATTGTT 312
Db 318 CTAATGATGCTGCTGGCGATGTACCACTTGTGCAACAGTACTCACCCGAGCAATATTG 377
QY 313 ATGAAGGACTAAAATGTGACAGCAGGTGCTTAATCCAATTTGTTATCCGTCGAGGAT 372
Db 378 CTGAAGGGTGCAGGCTGTGTAGCTGTGATGAATGCAATGGATCTTAGACAGAGGTATCA 437
QY 373 AAACAGCAACCAACAGCTGTTGAAGCCTTGAAGCCCTTGAAGCCCTTCAACCTGTATCTGGCA 432
Db 438 CAATGGCTGCTGATGCTGTTGTAACAACCTGAAAGCAAGACGCGATGATAGCACAT 497
QY 433 AGGAAGCTATTGCTCAGGTCGCTGCGAGTATCATCAGCTCTGAA---AAAGTTGGAGAGT 489
Db 498 CCGAGGAGATTGCCAGGTTGGTACAATCTCTGCAATGGAGAACGCGGAATAGGTGAGA 557
QY 490 ATATCTCAGAACTATGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGATCTC 549
Db 558 TAAATGCAAGGCTATGGAGCAAGTAGGAAAAGAGGTGTCTACATATTCAAGATGGAA 617
QY 550 GAGGTATGAAAACAGAACTTGAAGTGGTTGAAGGATGCAATTTGACCGTGTACCTGT 609
Db 618 AGACACTGCTTAATGATTTGCAAGTTGTTGAGGGGNTGAAGTTGACAGGGGCTACATAT 677
QY 610 CTCAATACATGTCACAGCAATGAAAATGTTGTCAGACCTTGAACACCCATTAT 667
Db 678 CCCCATACTTCATCACAATGAGAAGATCAGAAAATGTGAACCTGGATGACCCACTTAT 735
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:33:17 ; Search time 1836.42 Seconds  
(without alignments)  
3900.644 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661

Sequence: 1 GAATTCGGCTTCATATGGCA.....TGGCGGATAAGCCGAATTC 1661

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 12362702 seqs, 2156297445 residues

Total number of hits satisfying chosen parameters: 24725404

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/pna/pctus\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US097\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US099\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
51: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
52: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
53: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
54: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
55: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
56: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1661	100.0	1661	14	US-09-001-737-7 Sequence 7, Appli
2	1052	63.3	1654	14	US-09-001-737-3 Sequence 3, Appli
3	1039.2	62.6	1623	21	US-09-583-110-1024 Sequence 1024, Ap
4	1039.2	62.6	1623	42	US-60-174-089-100 Sequence 100, App
5	1038.4	62.5	1926	22	US-09-613-303-50 Sequence 50, Appli
6	1035.4	62.3	5365	13	US-08-961-527-77 Sequence 77, Appli
7	1024.4	61.7	3714	31	US-60-061-998-525 Sequence 525, App
8	1023.4	61.6	5373	27	US-60-029-960-223 Sequence 223, App
9	1016.6	61.2	2597	31	US-60-068-175-420 Sequence 420, App
10	945.8	56.9	8787	31	US-60-068-217-910 Sequence 910, App
11	945.2	56.9	8784	29	US-60-045-649-1072 Sequence 1072, Ap
12	945.2	56.9	8784	29	US-60-046-653-1063 Sequence 1063, Ap
13	942.6	56.7	3625	14	US-09-070-927-42 Sequence 42, Appli
14	942.6	56.7	3625	14	US-09-070-927A-42 Sequence 42, Appli
15	934	56.2	1638	15	US-09-134-000-1154 Sequence 1154, Ap
16	901.6	54.3	2237	31	US-60-068-186-529 Sequence 529, App
17	901.6	54.3	2300	30	US-60-050-444-525 Sequence 525, App
18	741.4	44.6	53996	12	US-08-831-156A-95 Sequence 95, Appli
19	741.4	44.6	278942	22	US-09-611-529-7437 Sequence 7437, Ap
20	735.6	44.3	1623	18	US-09-450-969-2422 Sequence 2422, Ap
21	704.2	42.4	1647	14	US-09-077-574A-1 Sequence 1, Appli
22	704.2	42.4	4972	14	US-09-077-574A-5 Sequence 5, Appli
23	701.6	42.2	6418	28	US-60-038-081-257 Sequence 257, App
24	701.6	42.2	6418	29	US-60-046-933-256 Sequence 256, App
25	701.6	42.2	6418	31	US-60-068-228-250 Sequence 260, App
26	657.4	39.6	1641	20	US-09-540-236-654 Sequence 654, App
27	657.4	39.6	1641	37	US-60-128-476-1558 Sequence 1558, Ap
28	654.2	39.4	96109	31	US-09-596-002-35 Sequence 35, Appli
29	654.2	39.4	96109	39	US-60-140-121-35 Sequence 35, Appli
30	652	39.3	1368	22	US-09-611-529-119 Sequence 119, App
31	649.6	39.1	24380	31	US-60-068-139-393 Sequence 393, App
32	649.6	39.1	24436	33	US-60-082-116-300 Sequence 300, App
33	624.8	37.6	1653	17	US-09-328-352-1382 Sequence 1382, Ap
34	606	36.5	1050	15	US-09-107-532-3030 Sequence 3030, Ap
35	590.6	35.6	1665	14	US-09-001-737-1 Sequence 1, Appli
36	589.8	35.5	1647	18	US-09-472-971-5 Sequence 5, Appli
37	589.8	35.5	4524	18	US-09-472-971-7 Sequence 7, Appli
38	572.8	34.5	1659	16	US-09-252-691-1518 Sequence 1518, Ap
39	572.8	34.5	1659	16	US-09-252-691C-1518 Sequence 1518, Ap
40	569.6	34.3	1674	16	US-09-252-691-1525 Sequence 1525, Ap
41	569.6	34.3	1674	16	US-09-252-691C-1525 Sequence 1525, Ap
42	568.6	34.2	966	55	US-09-107-433-1679 Sequence 1679, Ap
43	568.2	34.2	1662	14	US-09-001-737-5 Sequence 5, Appli
44	551.6	33.2	910715	1	PCT-US98-12764-1 Sequence 1, Appli
45	544.4	32.8	2416	16	US-09-221-017B-895 Sequence 895, App

# ALIGNMENTS

RESULT 1  
 US-09-001-737-7  
 : Sequence 7, Application US/09001737  
 : GENERAL INFORMATION:  
 : APPLICANT: Mizzen, Lee  
 : APPLICANT: Wisniewski, Jan  
 : TITLE OF INVENTION: STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE  
 : NUMBER OF INVENTIONS: HSP60 FAMILY  
 : NUMBER OF SEQUENCES: 91  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: SEED and BERRY LLP  
 : STREET: 701 Fifth Avenue, 6300 Columbia Center  
 : CITY: Seattle  
 : STATE: Washington  
 : COUNTRY: USA  
 : ZIP: 98104  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/001,737  
 : FILING DATE: 31-DEC-1997  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Sharkey, Richard G.  
 : REGISTRATION NUMBER: 32,629  
 : REFERENCE/DOCKET NUMBER: 870109.408  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (206) 622-4900  
 : TELEFAX: (206) 682-6031  
 : INFORMATION FOR SEQ ID NO: 7:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1661 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 15..1649  
 : US-09-001-737-7

Query Match 100.0%; Score 1661; DB 14; Length 1661;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCTTCATATGGCAAAAGAAATCAATTTTCACAGATGCGCGTGTGCCATGG 60  
 Db 1 GAATTCGGCTTCATATGGCAAAAGAAATCAATTTTCACAGATGCGCGTGTGCCATGG 60

QY 61 TCGCGGAGCTTGATATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTCAAGGCGCA 120  
 Db 61 TCGCGGAGCTTGATATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTCAAGGCGCA 120

QY 121 ATGTTGTTCTTGAAGAAAGCTTTTGGTTTCCCTTAATTAATGACGGGGTAAACCATTTG 180  
 Db 121 ATGTTGTTCTTGAAGAAAGCTTTTGGTTTCCCTTAATTAATGACGGGGTAAACCATTTG 180

QY 181 CTAAGAGATCGAATYTAGAAGATCAATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAG 240  
 Db 181 CTAAGAGATCGAATYTAGAAGATCAATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAG 240

QY 241 TGGCTTCTAAAACCAATGATATTTGGTGTGGGACGACTACTCAACAGTTTTCACAC 300  
 Db 241 TGGCTTCTAAAACCAATGATATTTGGTGTGGGACGACTACTCAACAGTTTTCACAC 300

QY 301 AAGCCATTTTTCATGAAGGACTAAAAAATGTGACACAGGTGCTAATCCAAATTTGGTATCC 360  
 Db 301 AAGCCATTTTTCATGAAGGACTAAAAAATGTGACACAGGTGCTAATCCAAATTTGGTATCC 360

QY 361 GTCGAGGCAATTTGAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAAC 420

Db 361 GTCGAGGCAATTTGAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAAC 420  
 QY 421 CTGTATCTGCGAAGGAAGCTATTGCTCAGTCCGCTCAGTATCATCACGCTCTGAAAGAG 480  
 Db 421 CTGTATCTGCGAAGGAAGCTATTGCTCAGTCCGCTCAGTATCATCACGCTCTGAAAGAG 480  
 QY 481 TTGGAGAGTATATCTCAGAAAGCTATTGAGCGGTGTGGGCAACGATGTTGATTAACCATCG 540  
 Db 481 TTGGAGAGTATATCTCAGAAAGCTATTGAGCGGTGTGGGCAACGATGTTGATTAACCATCG 540  
 QY 541 AAGATCTCGAGGTATGGAACAGAACTTTGAAGTGGTTTGAAGCATGCAATTTGACCGTG 600  
 Db 541 AAGATCTCGAGGTATGGAACAGAACTTTGAAGTGGTTTGAAGCATGCAATTTGACCGTG 600  
 QY 601 GTTACCTGCTCAATACATGTTGACAGCAAAATCAAAATTTGTTGAGACCTTTGAAACCC 660  
 Db 601 GTTACCTGCTCAATACATGTTGACAGCAAAATCAAAATTTGTTGAGACCTTTGAAACCC 660  
 QY 661 CATTTATCTTAATACAGGATATAAAGTGTCAAAACATCAAGACATTTTGCACACTACTTG 720  
 Db 661 CATTTATCTTAATACAGGATATAAAGTGTCAAAACATCAAGACATTTTGCACACTACTTG 720  
 QY 721 AGGAAGTCTTAAACCAACCGTCCATTACTCATTTATTCAGATGATGGATGGTGAAG 780  
 Db 721 AGGAAGTCTTAAACCAACCGTCCATTACTCATTTATTCAGATGATGGATGGTGAAG 780  
 QY 781 CACTTCCAAACCTTGTCTTGAACAGATTCGTTGATTTCAATGCTTCAATGCTTCAAG 840  
 Db 781 CACTTCCAAACCTTGTCTTGAACAGATTCGTTGATTTCAATGCTTCAATGCTTCAAG 840  
 QY 841 CGCCAGGATTTGGTGTATCGTCTGTAAGCTATGCTTGAAGACATTTGATCTTACAGAGTG 900  
 Db 841 CGCCAGGATTTGGTGTATCGTCTGTAAGCTATGCTTGAAGACATTTGATCTTACAGAGTG 900  
 QY 901 GTACAGTGAATACAGAGGATCTAGGACTTTGAATTTAAAGATGCTAACAATGACAGCCCTTG 960  
 Db 901 GTACAGTGAATACAGAGGATCTAGGACTTTGAATTTAAAGATGCTAACAATGACAGCCCTTG 960  
 QY 961 GACAGGCTCTAAGATTACAGTTGATAAGATAGACAGCTAATTTGTTGAAGTTTCAAGAA 1020  
 Db 961 GACAGGCTCTAAGATTACAGTTGATAAGATAGACAGCTAATTTGTTGAAGTTTCAAGAA 1020  
 QY 1021 GTTCAAGAGCTATTGCTAACCGTATTGCTACTGATTAATTAATGCAATTTAGAAACAACTT 1080  
 Db 1021 GTTCAAGAGCTATTGCTAACCGTATTGCTACTGATTAATTAATGCAATTTAGAAACAACTT 1080  
 QY 1081 CTGACTTTGACCGTGAAACAACTTACAGAAAGCTTTTGGCGAAATTTAGCTGCTGTTAGCTG 1140  
 Db 1081 CTGACTTTGACCGTGAAACAACTTACAGAAAGCTTTTGGCGAAATTTAGCTGCTGTTAGCTG 1140  
 QY 1141 TTATCAAGATAGGAGCTTCAACAGACAGCTTTTAAAGAAATGAACTTCCCATTTGAGG 1200  
 Db 1141 TTATCAAGATAGGAGCTTCAACAGACAGCTTTTAAAGAAATGAACTTCCCATTTGAGG 1200  
 QY 1201 ATGCTCTAAATGCTACACGTCAGCGCTTGAAGAGGATTCGTTGGTGGTGAAGACAG 1260  
 Db 1201 ATGCTCTAAATGCTACACGTCAGCGCTTGAAGAGGATTCGTTGGTGGTGAAGACAG 1260  
 QY 1261 CACTTATACGGTTATTGAAAAGTAGCAGCTTCTTGAAGCTTGAAGGCGATGATGCTACTG 1320  
 Db 1261 CACTTATACGGTTATTGAAAAGTAGCAGCTTCTTGAAGCTTGAAGGCGATGATGCTACTG 1320  
 QY 1321 GACGTAACATTTGCTGCTCTAGAGAGCGCTTCTAGCTCAAAATGCTTTAAATGCTG 1380  
 Db 1321 GACGTAACATTTGCTGCTCTAGAGAGCGCTTCTAGCTCAAAATGCTTTAAATGCTG 1380  
 QY 1381 GGTACGAAGGCTCCCTAGTATTGACAACTTGAAGAAACAGCCCTGCAGGAACAGGATTTA 1440  
 Db 1381 GGTACGAAGGCTCCCTAGTATTGACAACTTGAAGAAACAGCCCTGCAGGAACAGGATTTA 1440  
 QY 1441 ATGCTGCAACAGGCTGAGTGGGTTGATGATTAAGCAAGGAATCATTTGACCTGCTCAAG 1500

Db 1441 ATGCTGCAACAGGTGAGTGGTTGATATGATTAACACAGGAATCATTTGACCCCTGTCAAAG 1500  
QY 1501 TACACAGATCAGCGCTTCAAAATGACAGCTTCTGTAGTAGTCTTATTTTGAACACAGAG 1560  
Db 1501 TACACAGATCAGCGCTTCAAAATGACAGCTTCTGTAGTAGTCTTATTTTGAACACAGAG 1560  
QY 1561 CAGTGTGTTGTAATAACCTGACACAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620  
Db 1561 CAGTGTGTTGTAATAACCTGACACAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620  
QY 1621 ATCCAGGAATGATGGTGGGATGGCGGATGAAGCCGAATTC 1661  
Db 1621 ATCCAGGAATGATGGTGGGATGGCGGATGAAGCCGAATTC 1661

## RESULT 2

US-09-001-737-3  
: Sequence 3, Application US/09001737  
: GENERAL INFORMATION:  
: APPLICANT: Mizzen, Lee  
: APPLICANT: Wisniewski, Jan  
: TITLE OF INVENTION: STREPTOCOCAL HEAT SHOCK PROTEINS OF THE  
: TITLE OF INVENTION: HSP60 FAMILY  
: NUMBER OF SEQUENCES: 91  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SEED and BERRY LLP  
: STREET: 701 Fifth Avenue, 6300 Columbia Center  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: USA  
: ZIP: 98104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/001,737  
: FILING DATE: 31-DEC-1997  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 870109.408  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 622-4900  
: TELEFAX: (206) 682-6031  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1654 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 15..1637  
US-09-001-737-3

Query Match 63.3%; Score 1052; DB 14; Length 1654;  
Best Local Similarity 77.7%; Pred. No. 3e-273;  
Matches 1271; Conservative 0; Mismatches 365; Indels 0; Gaps 0;  
QY 1 GAATTCGGCTTCATATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGGCCATGG 60  
Db 1 GAATTCGGCTTCATATGGCAAAAGAAATTAATTTTCATCAGATGCCGCTTCAGCTATGG 60  
QY 61 TGGCCGGAGTTGATATGTTAGCAGATACCGTCAAGTAACGCTTGTGCTTAAAGGGCCCA 120  
Db 61 TCCGTGGTGTGATATCCTTGCAGACTGTTAAAGTAACCTTTGGGACCAAAAGGTGCGCA 120  
QY 121 ATGTTGTTCTTGAAGAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCAATTG 180  
Db 121 ATGTTGTTCTTGAAGAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCAATTG 180

Db 121 ATGTCGTTCTTGAAGAGCTCATTCGGTTACCCCTTGATTTACCAATGACGGTGTGACTATTG 180  
QY 181 CTAAGAGATCGAATTAGAAAGATCATTTTGAACACATGGGAGCAAAATTTGCTGTCTGAAG 240  
Db 181 CTAAGAGATCGAATTAGAAAGATCATTTTGAACACATGGGAGCAAAATTTGCTGTCTGAAG 240  
QY 241 TGGCTTCTAAACCAATGATATTCGTTGATGGGACGACTACTGTGCAACAGTTTGTGACAC 300  
Db 241 TAGCTTCAAAACCAATGATATTCGTTGATGGGAGCAAAATTTGCTGTCTGAACCTGCACTGTTTGAACC 300  
QY 301 AAGCCATTTGTTCAAGAGGACTTAAAGAAATGTGACAGCAGGTGCTTAATCCCAATTTGGTATCC 360  
Db 301 AAGCAATCGTCGTTGAAGGAAATCAAAACGTCACAGCAGGTGCAAAATCCCAATCGGTATTC 360  
QY 361 GTCGAGCATTTGAACACAGCAACAGCAGCTGTTGAAGCCTTTGAAAGCCATTTGCTCAAC 420  
Db 361 GTCGAGCATTTGAACACAGCAGCTGTTGAGCAGCTTTGAAAGCAACAGCTCATCC 420  
QY 421 CTGTATCTGGCAAGGAAGCTATTGCTCAGGTGCTGCTGAGTATCATACAGCTCTGCTGAAAAAG 480  
Db 421 CTGTTGCAATAAAGAGCTATCGTCAAGTTGACGCGGTATCTCTCGTTCTGAAAAAG 480  
QY 481 TTGAGAGATATCTCAGAGCTATGAGCGTGTGGCAACGATGCTGTGATTAACCATCG 540  
Db 481 TTGAGAGATATCTCAGAGCTATGAGCGTGTGGCAACGATGCTGTGATTAACCATCG 540  
QY 541 AAGAAATCTCGAGGTATGGAACAGAACTTGAAGTGGTTGAAGCATCAATTTGACCGTG 600  
Db 541 AAGATCACTGATGGAACAGAGCTTGAAGTGGTTGAAGCATCAATTTGACCGTG 600  
QY 601 GTTACCTGCTCAATCATGCTCAGACAAATGAAAAATGGTTGCGAGCTTGTGAAAAAC 660  
Db 601 GTTACCTTTCACAGTACATGCTGACAGATAGCGAAAAATGGTGGCTGACCTTGTAAAAATC 660  
QY 661 CATTATCTTAATCAGCGATTAAGAGGTGCAACATCCAGACATTTTCCACTACTTGG 720  
Db 661 CGTACATTTGATTTACAGCAAGAAATTTCCAAATATCCCAAGAAATTTGCGCACTTTGG 720  
QY 721 AGGAAGTTCTTAAACCAACCGCTCATTACTCATTTTGCAGATGATGTTGATGGTGAAG 780  
Db 721 AAGCATTTCTCAAGCAATCGTCCACTCTTGATTTGCGGATGATGTTGATGGTGAAG 780  
QY 781 CACTTCAACCCCTTGTCTTGAACAAAGATTCGTGCTACTTTCAATGCTGTTGCTGCTCAAG 840  
Db 781 CTCTTCAACTCTGTTTGAACAAAGATTCGTGGAACCTTCAACGTTAGTAGCAGCTCAAG 840  
QY 841 CGCCAGATTTGTTGATCGTCTGTAAGCTATGCTTGAAGACATTCCTTGTGACAGGTG 900  
Db 841 CACTGTTTGTGTCGCGCAAGCCATGCTTGAAGATATGCGCATCTTAAACAGGCG 900  
QY 901 GTACAGTATTACAGAGGATCTAGGACTTGAATTTAAAGATGCTACAAATGACAGCCCTTG 960  
Db 901 GAACAGTTATCAGAGAAGACCTTGGTCTTGAGTTGAAGATGCGCAATTTGAAGCTCTTG 960  
QY 961 GACAGGCTGCTAAGATTACAGTTGATTAAGATAGCAGCAATTTGTTGAAGTTTCAAGAA 1020  
Db 961 GTCAAGCAGCAGAGTGCACCGTGGCAAGATAGCAGCGTTATTGTTAGAGGTGCAAGAA 1020  
QY 1021 GTTCAGAGCTATTGCTAAACCGTATTCGACTGATTAATTCGCANTTAGAACAACACACTT 1080  
Db 1021 ATCCTGAAGCGATTTTCACCGGTGTTGCGGTATTCAGAGTCTCAAAATCGAACTACAAC 1080  
QY 1081 CTGACTTTGACCGTGAAGAACTACAAAGAGCTTTGGGCAAAATAGCTGGTGGTGTAGCTG 1140  
Db 1081 CTGATTTGACCGTGAAGAAATTCGAAGAGCTTTGGCAAAATTTGCTAGGTTGGTGTAGCG 1140  
QY 1141 TTATCAAAGTAGGAGCTTCCCAACAGAGACAGCTTTTAAAGAAATGAAATTTGCTGCTAGG 1200  
Db 1141 TTATTAAGGTGCGAGCGCAACTGAACTGAGTTGAAGAAATGAAATTCGCGATTGAAG 1200  
QY 1201 ATGCTCTAATGCTACAGTGCAGCGCTTGAAGAGTATGCTGCTGGTGGTGGGAACAG 1260  
Db 1201 ATGCCCTCAACGCTACTCGTGCAGCTGTTGAAGAGGTATTGTTGCAAGTGGTGGTGGGAACAG 1260

QY 1261 CACTTATTACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGCGGATGATGCTACTG 1320  
 Db 1261 CTCTTGCCAAATGTGATCCAGCTGTTGCTACCTTGGAAATTTGACAGGAGATGAAGCAACAG 1320  
 QY 1321 GAGCTACATTTGCTTCTGCTCTAGAGAGGCTGTAGTCAAAATGCTTTAAATGCTGTG 1380  
 Db 1321 GAGCTAAATTTGCTTCTGCTCTAGAGAGGCTGTAGTCAAAATGCTTTAAATGCTGTG 1380  
 QY 1381 GGTACGAAGGCTCCGTAGTTATTGCAAGTTGAAAAACAGCCCTGCAGGAACAGGATTGA 1440  
 Db 1381 GATTGGAAGATCTCGTTATCGATCGCTTTGAAAAATGCTGAGCTTGGTATAGGATCA 1440  
 QY 1441 ATGCTGCAACAGTGTAGTGGTGTATGATTAATAACAGGAATCATTTGACCCCTGTCAAAAG 1500  
 Db 1441 ACCGCAACTGCGAGTGGTGTATGATTAATAACAGGAATCATTTGACCCCTGTCAAAAG 1500  
 QY 1501 TACACGATCAGGCTTCAAAATGAGCTTCTGCTAGCTAGTCTTATTTGACAAACAGAG 1560  
 Db 1501 TGAGTCTGTTAGCCCTTACAAATGAGCTTCTGCTAGCTAGTCTTATTTGACAAACAGAG 1560  
 QY 1561 CAGTGTGCTTAATAACCTGAAACAGCTAGCCAGCCAGCAATGCCAGCAGGTATGG 1620  
 Db 1561 CAGTGTGCTTAATAACCTGAAACAGCTAGCCAGCCAGCAATGCCAGCAGGTATGG 1620  
 QY 1621 ATCCAGGAATGATGGG 1636  
 Db 1621 TGGTGGAAATGGCGG 1636

RESULT 3

US-09-583-110-1024  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al.  
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
 ; FILE REFERENCE: PATH00-07A  
 ; CURRENT APPLICATION NUMBER: US/09/583,110  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/107,433  
 ; PRIOR FILING DATE: 1998-06-30  
 ; PRIOR APPLICATION NUMBER: US 60/085,131  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: US 60/051,553  
 ; PRIOR FILING DATE: 1997-07-02  
 ; NUMBER OF SEQ ID NOS: 5322  
 ; SEQ ID NO 1024  
 ; LENGTH: 1623  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-583-110-1024

Query Match 62.6%; Score 1039.2; DB 21; Length 1623;  
 Best Local Similarity 77.6%; Pred. No. 8.7e-270;  
 Matches 1257; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

QY 15 ATGCGAAAAAGAAATCAAAATTTTACAGAGATCGCGTGGCTGCCATGGTGGCGGAGTGTAT 74  
 Db 1 atgcaaaagaaataaaattttcaacagatgcccgctcagccatggttcgtggtgcgat 60  
 QY 75 ATGTTAGCAGATACCGTCAAAATGCTTGGTCTTAAAGGGCGCAATGTTGCTTTGAA 134  
 Db 61 attcttgacacactgttaagtaaccttggacccaaaggtcgcgaatgtcttcttgaa 120  
 QY 135 AAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCATGCTTAAGAGATCGAA 194  
 Db 121 aagtcattcgttcaccccttgattaccatgacggtgtgaccatgccaagaagaatcgaa 180  
 QY 195 TTAGAGATCATTTTGAACAAATGGACAAATTTGGTCTCTCAAGTGGCTTCTAAACC 254  
 Db 181 ttggaagaccattttgaaaataatgggtgctaaagttagatcatcagaagtagcttctaaaacc 240

QY 255 AATGATATTGCTGGTATGGGACGACTACTGCAACAGTTTTTACACAAGCCATTGTTTCAT 314  
 Db 241 aatgataatcgaggtagcgaactacgactgcaacagttctgacccaagctatcgcctg 300  
 QY 315 GAAGGACTAAAAAATGTGACAGAGTGTAAATCCAAATTTGGTATCGTTCGAGCATTTGAA 374  
 Db 301 gaaggaaatcaaaaacgttcacagcaggtgcaaatcccaatcggtatttcgtcggtggattgaa 360  
 QY 375 ACAGCAACAGCAACAGCTGTTGAAGCCTTTGAAAGCCATTGCTCAACCTGTATCTGCAAG 434  
 Db 361 acagcagttgcccagcagttgaagctttgaaaaaacagccatcccttctgccaataaa 420  
 QY 435 GAAGCTATTGCTCAGCTCGCTGAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC 494  
 Db 421 gaagctatcgctcaagtgcagcagtatctctcgtcttgaaaaaagtgtgtgagtaac 480  
 QY 495 TCAGAAAGCTATGGAGCGTGTGGCAACAGATGTTGATTACCATCGAAGAAATCTCGAGGT 554  
 Db 481 tctgaagcaatggaaaaagttggcaaaagcgggtgcatcaccatcgaaagagtcacgtggt 540  
 QY 555 ATGGAACACAGAACTTCAAGTGGTTGAAGCATCCAAATTTGACCGTGGTTACCTGTCTCAA 614  
 Db 541 atggaacacagacttgaagtcgtagaagaaatgcagtttgaccgtggtcacctttccacag 600  
 QY 615 TACATGGTCACAGCAATGAAAAAATGGTTGCAGACCTTTGAAAAACCCATTATCTTAATC 674  
 Db 601 tacatggtgacagatagcgaataatggtggtgacacttgaaaaatccgtacattttgatt 660  
 QY 675 ACGGATAAAAAAGTGTCAAAACATCCAAAGACATTTTGCCACTACTTGAGGAAGTTCCTTAA 734  
 Db 661 acagacaagaataatcccaacatcccaagaatacttgccacttttggaagacattctccaa 720  
 QY 735 ACCAACCTGCCATTACTCATTTATTCAGATGATGTGGATGGTTCAGACACTTCCAACCTTT 794  
 Db 721 agcaatcgctccactcttgatttcggatgatgtggtggtgagctcttccactctt 780  
 QY 795 GTCTTGAACAAGATCGTGGTACTTTCAATGTGGTGTGCTCAAAAGCGCAGGATTTGTT 854  
 Db 781 gttttgaacaagatcgttgaaaccttcaacgttagtagcagtcgaagcaacctggttttgg 840  
 QY 855 GATCGTGTAAAGCTATGCTTGAAGACATTTGCTATCTTGACAGGTGGTACAGTATTACA 914  
 Db 841 gaccgtgcgaagccatgcttaagataatcgcaatcttaacagcggaacagcttatcaca 900  
 QY 915 GAGGATCTAGGACTTCAATTAAGATGCTACAATCAGACAGCCCTTGCGAGAGTCTCTAG 974  
 Db 901 gaagaccttggctcttgatgaaagatgcgaataatgaaagctcttggccaagcagcaga 960  
 QY 975 ATTACAGTTGATAAAGATAGCACAGTAAATTTGAAGGTTTCAGGAAGTTTCAGAAAGTATT 1034  
 Db 961 gtgacctggacaagatagcagcgttattgtagaaggtgcaggaatactcctgaagcatt 1020  
 QY 1035 GCTAACCGTATTGCACTGATTAATTCGCAATTTAGAAACAACTTCTGACTTTGACCGT 1094  
 Db 1021 tctccacgtgttgctggttatcaagtctcaaatcgaaactacaaactctgaaatttgacct 1080  
 QY 1095 GAAAACTACAGAAGCTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTTATCAAGTAGGA 1154  
 Db 1081 gaaaaatgcaagaacgccttgcccataattgtcaggtggtgtagcgttattaaagtcgga 1140  
 QY 1155 GCTCCAACAGACAGCTTTTAAAGAAATGAACTTCGATTGAGGATGCTCTAAATGCT 1214  
 Db 1141 gccgcaactgaaactgagttgaaagaaatgaaactccgcattgaaagatgccctcaacgt 1200  
 QY 1215 ACAGTGCAGCGGTTGAAGAAGGTATCGTGTGCTGGTGGGAACACACTTATTACGGTT 1274  
 Db 1201 actcgtgcagctgttgaagaaggtattgttcaggtggtggaaacagctcttgcgaatgtg 1260  
 QY 1275 ATTGAAAAAGTAGCAGCTCTTACAGCTTGGGCGGATGATGCTTACTTGGAGCTTAACATTG 1334  
 Db 1261 attccagctgtgttcaccttgaattgacagagatgaaagcaacaggaagcgttaattgtt 1320









```

Db 1657 TATCGTTATGATCGTTTGAATAAATGCTGAGCTGTGATAGGATTTAAACGACGAACTGG 1716
QY 1454 TGAGTGGGTTGATATGATTTAAACAGGAATCAATGACCCCTGTCAAAAGTAAACAGATCAGC 1513
Db 1717 CGAGTGGGTTACATGATGATCAAGTATCATGATCCAGTTAAAGTGAGTGGTTCAGC 1776
QY 1514 GCTTCAAAATGAGCTTCTGTAGCTAGTCTTATTTTGAACAAGAGCAGTGTGCTAA 1573
Db 1777 CCTACAAAATGAGCATCTGTAGCCAGCTTGATTTTGAACAAGAGCAGTGTGAGCCAA 1836
QY 1574 TAAACCTGAACAGCTACCCAGCCAGCCAGCATGCCAGCAGGTATGATCCAGGAATGAT 1633
Db 1837 TAAACAGAACAGTAGTACGCCAGCTCCAGCAATGGATCCAAAGCATGATGGCGGGATGAT 1896
QY 1634 G 1634
Db 1897 G 1897

RESULT 7
US-60-061-998-525
: Sequence 525, Application US/60061998
: GENERAL INFORMATION:
: APPLICANT: LAGACE, ROBERT E.
: APPLICANT: CORLEY, NEIL C.
: APPLICANT: RUSSO, FRANK D.
: APPLICANT: HANN, AMY L.
: APPLICANT: HEATH, JOE D.
: APPLICANT: FINNEY, GREGORY L.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
: TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
: NUMBER OF SEQUENCES: 797
: CORRESPONDENCE ADDRESS:
: ADDRESS: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/061,998
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: CERRONE, MICHAEL C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PM-0006-2P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 525:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3714 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: SPN1C530
US-60-061-998-525

Query Match 61.7%; Score 1024.4; DB 31; Length 3714;
Best Local Similarity 77.4%; Pred. No. 1.2e-265;
Matches 1255; Conservative Indels 1; Gaps 1;
QY 14 TATGCCAAAGAAATCAAAATTTTACAGCATGCGCGTGTGCCATGCGCGGAGTTGA 73

```

```

Db 224 TATGTCAAAAGAAATTAATTTTTCATCAGATGCCCGTTTCAGCCATGTTTCGTGGTTCGA 283
QY 74 TATGTTAGCAGATACCGTCAAAAGTAAAGTTCCTTAAAGGCGCAATGTTCTTCTTGA 133
Db 284 TATCCTTGCAGACACTGTTTAAAGTAACTTGGGACCAAAAGGTGCGCAATGTCGTTCTGA 343
QY 134 AAAAGCTTTTGTCTCTCCCTTAATTAATGACGGGGTAAACCATTTGCTAAAGAGATCGA 193
Db 344 AAAGTCATTCGGTTTACCCCTTGATTACCAATGACGCTGTGACCATTTGCCAAAGAAATCGA 403
QY 194 ATTAGAAGATCATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAAGTGGCTTCTTAAAC 253
Db 404 ATTGGAAGACCATTTTGAATAATATGGTGTCTAAGTTAGTATCAGAAGTAGCTTCTTAAAC 463
QY 254 CAATGATATTGCTGTGATGGGACGAC-TACTGCAACACATTTTGGACACAGCATTTGTTTC 312
Db 464 CAATGATATCGCAGGTGACGGGACTACGGACTGCAACAGCTTTGACCCCAAGTATATCGTCC 523
QY 313 ATGAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTTGGTATCCGTCGAGGCATTG 372
Db 524 GTGAAGGAATCAAAACGTCACAGCAGGTGCAAAATCCAATCGGTATTCGTCTGGTGGATTG 593
QY 373 AAACAGCAACAGACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGCGGA 432
Db 584 AAACAGCAGTTGCCGACGAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATA 643
QY 433 AGSAGCTATTGCTCAGGTGCGTGCAGTATCATCAGCTCTGAAAAAGTTGGAGAGTATA 492
Db 644 AAGAAGCTATCGCTCAAGTTGACGCGGTATCTCTCGTTCTGAAAAAGTTGGTGGAGTACA 703
QY 493 TCTCAAGACTATGAGCAGCTGTGGGCAACGATGGTGTGATTACCATCGAAGATCTCGAG 552
Db 704 TCTCTGAAGCAATGGAAAAAGTTGGCAAGACGGTGTCTATCACCATCGAAGAGTCAAGTG 763
QY 553 GTATGGAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTTC 612
Db 764 GTATGGAACAGAGCTTGAAGTGGTGAAGAGGAATGCAAGTTGACCGTGGTTACCTTTTAC 823
QY 613 AATACATGTCACAGACAATGAAAAAATGTTGCGAGACCTTGAACCCCATTTATCTTAA 672
Db 824 AGTACATGGTGACTGATACGGAAAAAATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 883
QY 673 TCACGGATAAAAAAGTGTCAAAACATCCAAAGACATTTTGGCAGTACTTTGAGGAAGTTCTTA 732
Db 884 TTACAGACAAGAAATTTCCATATCCAAAGAAATCTTGCCACTTTTGGAAAGCATTTCTCC 943
QY 733 AAACCAACCGTCCATTACTCATTTATTTGAGATGATGTGGATGGTGAAGCAGTTCACACCC 792
Db 944 AAAGCAATCGTCCACTCTTGATTATTTGCGGATGATGTGGATGGCGAGGCTCTTCCAACTC 1003
QY 793 TTGCTCTTGAACAAGATTCGTGCTACTTTTCAATGTGGTTCCTCTCAAGCGCCAGGATTTG 852
Db 1004 TTGTTTGAACAAGATTCGTGGAACCTTCAACGTAGTAGCAGTCAAGGACCTCTGTTGTTG 1063
QY 853 GTGATCGTGTAAAGCATGCTTTGAAGACATTCGTATCTTTGACAGGTGTACAGTATTA 912
Db 1064 GTGACCGTGCAGAACCCATGCTTTGAAGATATCGCCATCTTACAGCGGCAACATATCA 1123
QY 913 CAGAGATCTAGGACTTGAATTTAAAGATGCTTACATGACAGCCCTTGGACAGGCTGCTA 972
Db 1124 CAGAAGACCTTGTGCTTGAAGTGAAGATGCGACAATTTGAAGCTCTTGGTCAAGCAGCA 1183
QY 973 AGATTACAGTTTGAAGATGACAGTAATTTGTAAGGTTTCAGGAAGTTCAGAGCTA 1032
Db 1184 GAGTACCGTGGACAAGATAGCAGGTTTATTTGTAAGAGGTGCAGGAATCTCTGAAGCGA 1243
QY 1033 TTGCTAAACCGTATTGCAGCTGATTAATGCAATTTAGAAACAACACTTCTGACTTTGACC 1092
Db 1244 TTTCTCACCGTGTGGGTTATCAAGTCTCAATCGAAGACTACAACTTCTGAAATTTGACC 1303
QY 1093 GTGAAAAACTACAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGTGTGTTATCAAGTAG 1152

```

Db 1304 GTCAAAATTTGCAAGACCGCTTGGCCAAATGTGCAGTGTGTAGCGGTTATTAAAGTTG 1363  
QY 1153 GAGCTCCACAGACAGACGCTTTAAAGAAATGAATCTCGCATTTAGGATGCTCTAAATG 1212  
Db 1364 GAGCCGCAACTGAACCTAGGTTGAAAGAAATGAATCTCGCATTTAGGATGCTCTCAACG 1423  
QY 1213 CTACAGCTGCAGCGGTTGAAGAGGTTATCGTTGCTGGTGGTGAACACGACTTATTACGG 1272  
Db 1424 CTACTCTGCGAGCTGTTGAAGAAAGTATTGTTGACAGTGGTGAACAGCTCTTTGCCAATG 1483  
QY 1273 TTATTGAAAGATGAGCAGCTCTTGAAGCTTGAGGCGGATGATGCTACTTGGAGCTAAACATTG 1332  
Db 1484 TGATTCAGCTGTTGCTACCTTGGAAATTTGACAGAGATGAAGCAACAGGACGTAATATTG 1543  
QY 1333 TGCTTCGCTCTAGAGAGCCTGTACGTCAAAATTTGCTTTAAATGCTGGGTACGAAGGCT 1392  
Db 1544 TTCTCGTGCTTTGGAAGAACCGCTTCGTCAAAATTTGCTCAAAATTTGAGGAT 1603  
QY 1393 CCGTAGTTATTGACAAAGTTGAAGAAACAGCCCTCCAGGAACAGAGATTTAAATGCTGCAACAG 1452  
Db 1604 CTATCGTTATCGATCGTTTGAAGAAATTTGCTGAGCTTGGTATAGGATTTAAGCGACCACTG 1663  
QY 1453 GTGAGTGGGTTGATGATTTAAACAGGAATCATTTGACCCCTGTCAAGATGAACACGATCAG 1512  
Db 1664 GCAGTGGGTTAATGATTTGATCAAGGTATCATTTGATCCAGTTAAAGTGAAGTGGTTGCTAG 1723  
QY 1513 CGCTTCAAAATGAGCTTCTGTAGCTAGTCTTATTTTGAACAGAGAGAGTGGTTGCTA 1572  
Db 1724 CCTCAAAATGAGCATCTGTAGCCAGCTTGTATTTGACACAGAGAGAGTGGTACGCA 1783  
QY 1573 ATAACTGAAACAGCTACGCCAGCCAGCCAGCAATGCGCAGAGTATGGATCCAGGAATGA 1632  
Db 1784 ATAAAGCAGAACAGTAGAGCCAGCTCCAGCAATGGATCCAGCATGATGGCGGGATGA 1843  
QY 1633 TG 1634  
Db 1844 TG 1845

RESULT 8

US-60-029-960-223  
; Sequence 223, Application US/60029960  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 1649  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/60/029,960  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340PP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 223:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5373 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-60-029-960-223  
  
Query Match 61.6%; Score 1023.4; DB 27; Length 5373;  
Best Local Similarity 77.4%; Pred. No. 2,7e-265;  
Matches 1254; Conservative 0; Mismatches 366; Indels 1; Gaps 1;  
  
QY 14 TATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGTGCCATGTGCGCGAGTTGA 73  
Db 272 TATGTCAAAAGAAATTAATTTTCATCAGATGCGCGTGTGCCATGTGCGTGTGTCGA 331  
QY 74 TATGTTAGCAGATACCGTCAAGTAAAGTAAAGGCGGCAATGCTGCTCTTGA 133  
Db 332 TATCCTTTGCAGACACTGTTAAAGTAACTTTGGGACCAAAAGGTGCGCAATGTCGTTCTTGA 391  
QY 134 AAAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGTAAACCATTTGCTAAAGAGATCGA 193  
Db 392 AAAGTCATTCGGTTCACCTTTGATTAACCAATGACGGTGTGCCATTTGCCAAGAAATCGA 451  
QY 194 ATTAGAAGATCATTTTGAAGCAATGGGAGCAAAATTTGGTGTCTGAAAGTGGCTTTTAAAC 253  
Db 452 ATTGGAAGACCAATTTTGAAGATATGGGTGCTAAGTTAGTATCAGAAGTAGCTTTCTAAAC 511  
QY 254 CAATGATATTTGCTGGTATGCGGAGGACTACTGCAACAGTTTTTGACACAAAGCCATTGTCA 313  
Db 512 CAATGATATTCGAGGTTGACGGAATCTAGACTGCAACAGTCTTGACCCAAAGCTATCGTCCG 571  
QY 314 TGAAGGACTAAAAATGTGACAGCAGGTGCTAATCCAAATTTGGTATCGTGCAGGCAATTGA 373  
Db 572 TGAAGNATCAAAACGTCACAGCAGGTGCAATCCAAATCGGTATTCGTGCGGATGA 631  
QY 374 AACAGCAACGAAACAGCTTTGAAAGCCTTTGAAAGCCTTTGCAACCTTTGCTCAACCTGTATCTGCA 433  
Db 632 AACAGCAGTTGCGCGCAGCAGTTGAAAGCTTTGAAAGCAACAGCCCTCTGTTGCCAATAA 691  
QY 434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCAGCTCTGAAAAAGTTGGAGGATATAT 493  
Db 692 AGAAGCTATCGCTCAAGTTGCGAGCCGTTCTCTCGTTCTGAAAAAGTTGGTGAATACAT 751  
QY 494 CTCAGAGCTATGAGGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGATCTCGAGG 553  
Db 752 CTCTGAGCAATGGAAGAGTTGGCAAGAGCGGTGTCATCACCATCGAAGANGCACGTGG 811  
QY 554 TATGGAACAGAACTTTGAAGTGGTTGAAGCAGTGCATTTGACCGTGGTTACCTGTCTCA 613  
Db 812 TATGGAACAGAGCTTTGAAGTGGTAGAAGGAATGCAGTTTGAACGCTTTTACCTTTTCA 871  
QY 614 ATACATGGTCAACAGCAATGAAAAATTTGTCAGACCTTTGAAAAACCCATTATCTTAAT 673  
Db 872 GTACATGGTGCAGATAGCGAAAAATTTGGGCTGAGCTTTGAAAAATCCGTTACATTTTAT 931  
QY 674 CACGGATAAAAAAGTGTCAAAACATTTCCCACTACTTCCCACTACTTCCAGGAGTTCTTAA 733  
Db 932 TACAGACAGAAATTTCCCAATATCCAAAGAAATCTTGCCACTTTTGGAAAGCATTTCTCA 991  
QY 734 AACCAACCGTCCATTACTCATTTATTCAGATGATGTGGATGGTGAAGCACTTCCAAACCT 793  
Db 992 AAGCAATCGTCCACTCTTGATTTATGGGATGATGTGGATGGGAGGCTCTTCCAACCTCT 1051  
QY 794 TGCTTTGAACAGATTCGTGGTACTTTCAATGTGGTGTGCTCAAGCGCCAGGATTTGG 853  
Db 1052 TGTTTTGAACAGATTCGTGGAACTTTCAACGTA-TAGCAGTCAAGGACACCTGGTTTGG 1110  
QY 854 TGATCGTGTAAAGCTATGCTTGAAGACATTTGCTATCTTGCAGAGTGTGACAGTATAC 913  
Db 1111 TGACCGTGCAGAAAGCCATGCTTGAAGATATCGCCATCTTAAAGCGGAGACATTTATC 1170  
QY 914 AGAGGATCTAGGACTTTGAATTTAAAGATGTACAATGACAGCCCTTTGGACAGGCTGCTAA 973

Db	1171	AGAAGACCTTGGTCTTGAAGTTCGACAAATGAAAGCTCTTGGTCAAGCAGCGAG	1230
Qy	974	GATTACCGTTGATAAAGATAGCACAGTAATTTGTTGAAGGTTTCAGGAAGTTTCAGAAGCTAT	1033
Db	1231	AGTGACCGTGGACAAAGATAGCACGGTTATTGTGAAGTTCGAGAAATCCTGAAGCGAT	1290
Qy	1034	TGCTAACCGTATTGCATGATTAAATTCGAATTGAACAACAACCTCTGACCTTTGACCG	1093
Db	1291	TTCTCACCGTGTGGGTTATTCAAGTCTCAAAATCGAACTACAACCTCTGAAATTTGACCG	1350
Qy	1094	TGAAAACCTACAGACAGCTTTGCGCAAAATAGCTGGTGGTCTAGCTCTTATCAAGTAGG	1153
Db	1351	TGAAAATTTGCAGACAGCTTTGGCCAAATTTGTAGTGGTGTAGCGGTTATTAAAGGTTGG	1410
Qy	1154	AGCTCCAACGAGACAGCTTTTAAAAAAGTAATGAAACTTCGCAATTCAGGATGCTCTAAATGC	1213
Db	1411	AGCGCAACTGAAACTGAGTTTGAAGAAATGAAATCCGCATTGAAGATGCCCTCAACGC	1470
Qy	1214	TACACGTGCAGCGTTGAAGAGGTATCGTTGCTGGTGGTGGACACAGCACTTATTACGGT	1273
Db	1471	TACTCGTGCAGCTGTTGAAGAAAGTATTGTTGCAAGTGGTGGAAACAGCTCTTGCCAATGT	1530
Qy	1274	TATTGAAAAGTAGCAGCTCTTGAGCTTTGAGGGCGATGATCTACTCGACGTAACATTGT	1333
Db	1531	GATTCAGCTGTGCTACCTTGGAAATTCACAGGAGATGAACACACAGCAGTATATTGT	1590
Qy	1334	GCCTCGTCTCTAGAAGAGCCCTGTACGTCAAATTTGCTTTAAATGCTGGGTACGAAGCTC	1393
Db	1591	TCCTCCGTGTTTGGGAAGAACCCGTTGCTCAAATTTGCTCACAAATCGAGGATTTGAAGGATC	1650
Qy	1394	CGTAGTTATTGACAGTTTGAAAACAGCCCTCGCAGGACAGGATTTAATGCTGCAACAGG	1453
Db	1651	TATCGTTTATCGGATCGTTTGAAAATTTGCTGAGCTTGGTATAGGATTTTAAACCGCAACTGG	1710
Qy	1454	TGAGTGGGTTGATATGATTAAACAGGAATCATTTGACCCGTCAAACTGAACAGTACAGTACG	1513
Db	1711	CGAGTGGGTTACATGATTTGATCAGGTTATCATTTGATCCAGTTTAAAGTGAGTGGTTCAGC	1770
Qy	1514	GCCTCAAATGCGAGCTCTGTAGCTAGTCTTTATTTTGACACAGAACAGTCTGTGCTAA	1573
Db	1771	CCACAAAATGAGCATCTGTAGCCACGCTTGATTTTGACACAGAACAGTCTGTAGCCAA	1830
Qy	1574	TAACTCTGAACAGCTACGCCAGCGCCAGCAATCCGACAGGTTATGATCCAGGAATGAT	1633
Db	1831	TAAACCAAGCAAGCTAGCCCCAGCTCCAGCAATGGATCCAGCATGATGGCGGGGATGAT	1890
Qy	1634	G 1634	
Db	1891	G 1891	

RESULT 9  
US-60-068-175-420

```

; Sequence 420, Appication US/60068175
;
; GENERAL INFORMATION:
;
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
;
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
;
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
;
; NUMBER OF SEQUENCES: 1175
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;
; STREET: 3174 PORTER DRIVE
;
; CITY: PALO ALTO
;
; STATE: CALIFORNIA
;
; COUNTRY: USA
;
; ZIP: 94304
;
; COMPUTER READABLE FORM:
;

```

```
Db 830 AGTACATGTCACAGATAGCGCAAAAATGCTGGCTGACCTGAAATCCGTACATTTGA 889
Qy 673 TCACGGATAAAAAGTGTCAACATCCACAGACATTTTGCACACTACTTGAGAAAGTTCTTA 732
Db 890 TTACAGACAAAATTTTCCAATATCAAGAGATCTTGCCACTTTTGGAAAGCATCTCC 949
Qy 733 AAACCAACCGTCCATTTACTCATTTATTCACAGATGATGGATGAGTGAAGCACTTCCAAACC 792
Db 950 AAAGCAATCGTCCACTCTTGATTATTTGCGGATGATGGATGCGGAGGCTTCTTCCAAC 1009
Qy 793 TTGCTTTGAACAAGATTGCTGGTACTTTCAATGTTG-CTGTCAAAGCGCCAGGATTT 851
Db 1010 TTGTTTGAACAAGATCTGGGAACCTTCAACGTAGTACAGTCAAGCACTTGGTTT 1069
Qy 852 GGTGATCGTGTGAAGCTATGCTTGAAGACATTCATCTTTGACAGTGGTGTACAGTGATT 911
Db 1070 GGTGACCGTGCAGAACCATCTGTTGAAGATATCGCCATCTTAACAGCGGAACATTATC 1129
Qy 912 ACAGAGATCTAGGACTTGAATTAAGAGATGCTACATGACAGCCCTTGGACAGGCTGCT 971
Db 1130 ACAGAGACCTTGGTCTTGAGTTGAAGATGCGACAATTAAGACTCTTGTCAAGCAGCG 1189
Qy 972 AAGATTACAGTTGATAAAGATACACAGTAATTTGTTGAAGTTTCAGGAAGTTCAGAAAGCT 1031
Db 1190 AGAGTACCGTGCAGAAAGATACACCGTTATTTGTGAAGGTCAGGAATCTCGAAGCG 1249
Qy 1032 ATTGCTAACCGTATGTCACATGATTAATCGCAATTAAGAAACAACAACCTTCTGACTTTGAC 1091
Db 1250 ATTTCTACCGCTGTGCGGTTATCAAGTCTCAAAATCGAATCAAACTTCTGAATTGAC 1309
Qy 1092 CGTGAACAACTACAGAACGTTTGGCGAATTAAGCTGGTGGTGTAGCTGTTATCAAGTA 1151
Db 1310 CGTGAACAAATGCAAGAACGTTTGGCAAAATTTGTCAAGTGGTGTAGCGGTTATTAAGGTC 1369
Qy 1152 GGAGCTCCACAGACAGCTTTAAAGAAATGAAATTCGCAATTCGCAATTCAGGATGCTTAAT 1211
Db 1370 GGAGCGCAACTGAACTGAGTTGAAGAAATGAAATTCGCAATTCGCAATTCAGGATGCTTAAT 1429
Qy 1212 GCTACAGCTGACCGCTTGAAGAGTATCGTTGCTGGTGGTGAACAGCACTTATTACG 1271
Db 1430 GCTACTCGTCAGCTGTTGAAGAGTATTTGTCAGCGCGTGAACAGCTCTTGCAAT 1489
Qy 1272 GTTATTGAAGAAAGTACAGCTCTTGAGCTTGAGGGGATGATGCTACTGACGCTAAT 1331
Db 1490 GTGATTCAGCTGTGCTACCTTGGAAATTCAGAGGATGAAGCAACAGGACGTAAT 1549
Qy 1332 GTGCTTCGCTCTAGAAGAGCTGTACGTCAAATGCTTTAAATGCTGGGTACGAAGGC 1391
Db 1550 GTTCTCGGCTTGAAGAACCGCTGCTCAAAATGCTCAAAATGCAAGTGAAGGATGAAGGA 1609
Qy 1392 TCCGTAGTTATGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACA 1451
Db 1610 TCTATCGTTATCGATCGTTTGAATAATGCTGAGCTTGATAGGATTTAAGCGAGCACT 1669
Qy 1452 GCTGAGTGGTGTATGATGATTAACAGAGGATCAATGACCCCTGTCAAAGTAAACAGCATCA 1511
Db 1670 GCGAGTGGTGTAAACATGATGATGATCAAGGTATCATTTGATCCAGTTAAAGTGAAGTCA 1729
Qy 1512 GCGCTCAAAATGACGCTCTGTAGTGTAGTCTTATTTTGAACAACAGAGCAGTGTGTGCT 1571
Db 1730 GCGCTCAAAATGACGATCTGTAGCCAGCTTGTATTTTGAACAACAGAGCAGTGTAGCC 1789
Qy 1572 AATAACCTGAACAGCTACGCCAGCGCCAGCAATGCCAGGATGATGATCCAGGAATG 1631
Db 1790 AATAACCAAGAACAGTAGCCCGAGCTCCAGCAATGATGATGATGATGATGATGATGATG 1849
Qy 1632 ATG 1634
Db 1850 ATG 1852
```

RESULT 10  
US-60-068-217-910

```
; Sequence 910, Application US/60068217
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0003-5 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 910:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: EFALC910
; US-60-068-217-910

Query Match 56.9%; Score 945.8; DB 31; Length 8787;
Best Local Similarity 74.1%; Pred. No. 3.2e-244;
Matches 1196; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

Qy 15 ATGCGAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGTGCTGCGCGGAGTTGAT 74
Db 403 ATGCGAAAAGAGATTAATTTGCGAGAGATGCGAGTCAGCAATGCTACGCGGAGTAGAT 462
Qy 75 ATGTTAGCAGATCCGTCAAAGTAACCGTTGGTCTTAAAGGGCGCAATGTTGTTCTTGA 134
Db 463 GTATTAGCAGATACAGTGAAGTGACATTAGGCGCTTAAAGTCTGTAACGTTGTTTGA 522
Qy 135 AAGCTTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 194
Db 523 AATCATTTTGGTTCACCATTTGATTACTTAAGCATGGGGTAAACAATTTGTAAGGAAAT 582
Qy 195 TTAGAAGATCATTTTGAACATGCGGAGCAAAATTTGCTGTGAAGTGGCTTCTTAAAC 254
Db 583 TTAGAAGATCATTTTGAACATGCGGAGCAAAATTTGCTGTGAAGTGGCTTCTTAAAC 642
Qy 255 AATGATATTGCTGGTGTGAGGACGACTACTGCAACAGTGTGACACAGCAATTTGTT 314
Db 643 AATGATATTGCTGGTGTGAGGACGACTACTGCAACAGTGTGACACAGCAATTTGTT 702
Qy 315 GAAGGACTAAAAATGTGACAGCAGGCTGCTAATCCCAATTTGGTATCCGTCGAGGCA 374
Db 703 GAAGGCTTAAAAACGTAACGCTGCTGGAGCTTAACCCATTAGGTATTCGCCGTTGGA 762
```





Db	523	AAATCATTTGGTTCACCATGATTACTAACGATGGGGTAACAATTGCTATAGGAAATTTGAA	582
Qy	195	TTAGAAGATCATTTTGAACAACATGGAGACAAATTTGGTCTGAAGTGGCTTCTTAAACCC	254
Db	583	TTAGAAGATCATTTTGAACAACATGGCGCAAAATTTAGTTTTCAGAAGTTGCTTCTTAAACG	642
Qy	255	AATGATATTGCTGGTGATGGGAGCACTACTGCAACAGATTTTGACACAAGCCATTTGTTTCAAT	314
Db	643	AATGATATTGCTGGTGACGGAAACAACACAGCGACATGTTTTCAGACAAGCCATTTGTTTCTG	702
Qy	315	GAAGGACTTAAAAAATGTGACACAGGTGCTAATCCAAATTTGGTATCCGTTCGAGGCATTTGAA	374
Db	703	GAAGGCTTAAAAACGTACTGCTGGAGCTAACCCRTTAGGTATTCGCCGTGGGATTTGAA	762
Qy	375	ACAGCAACAGCAACAGCTGTTGAAGCCTTTGAAGCCATTGCTCAACCTGATCTGGCAAG	434
Db	763	TTAGCAACAAAAACACAGCTAGAAGAAATTACACAATAATTCATCTGTAGTTGATTTCAAAA	822
Qy	435	GAAGCTATTGCTCAGGTGCGTCAGTATCACCGCTCTGAAAAAGTTGGAGAGTATATTC	494
Db	823	GAAGCGATTGCACAAGTCGCTGTTTCATCAGCTTCTGAAAAAGTCGGCAATTAATTT	882
Qy	495	TCAGAAGCTATGAGCGTGTGGGCAACGATGCTGTCATTCACATCGAGAATCTCGAGGT	554
Db	883	GCCGATGCAATTGAAAAAGTTGGTACGACGCGGTAATTCACATTGAAGAAATCAAAAGG	942
Qy	555	ATGAAAAACAGAACTTCAAGTGGTTGAAGCATGCAATTTGACCGTGGTTACCTGCTCTCAA	614
Db	943	ATTGAACAGAAATTAGATGGGTTGAGGAATGCAATTTTCGACCGCGGTTATTATCTCAA	1002
Qy	615	TACATGTCACAGACAATGAAAAAATGGTTGCGACCTTTGAAAAACCCATTTATCTTAAATC	674
Db	1003	TACATGGTTACTGACAACGATAAAATGGAAGCTGTTTGTAGAAATCCATATATCTTAAAT	1062
Qy	675	ACGGATAAAAAGCTTCAACATCCAAGACATTTGCCACTACTTGAGGAAGTTCTTAAA	734
Db	1063	ACCGAAAAAAATCTCAAATTTCAAGAATCTTACCTTTATTAGAACAAATTTCTACAA	1122
Qy	735	ACCAACCGTCCATTACTCATTTATTCAGATGATGGATGGTGAAGACATTTCCAACCCCTT	794
Db	1123	CAAAGCGTCCACTATTGATTATTCGGGATGATGTTGATGGGAAGCTCTACCAACATTTA	1182
Qy	795	GTCPTGAACAAGATTCGTGGTACTTTCAATGGGTTGCTGCTAAAGCGCCAGGATTTGGT	854
Db	1183	GTATTGAACAAAAATCCGTGGTACATTTAATGTTGCGCAGTAAAGCGCCAGGATTTGGT	1242
Qy	855	GATCGTCGTTAAAGCTATGCTTGAAGACATTTGCTATCTTTGACAGGTGGTACAGTATTACA	914
Db	1243	GACGCGCAAAAGCGATGCTTGAAGATATTGCTATTATTAAACAGGTGGTACAGTAATCACT	1302
Qy	915	GAGGATCTAGGACTTGAATTTAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTTAAG	974
Db	1303	GACGCTTTAGGGTTAGAGTTTAAAGACACAACTATTGAAAACCTTAGGAAATGCTTAGCAAA	1362
Qy	975	ATTACAGTTGATAAAGATAGCACGATAATTGTTGAAGTTTCAGGAAGTTTCAGAAAGCTATT	1034
Db	1363	GTAGTTCTGCACAACAGATAACACACAATTTGCGAAGTGGCTGGTTCAAAGAAGCCATT	1422
Qy	1035	GCTAACCGTATTGCACTGATTAAATGCGAAATTAGAAAAACAACAACTCTCTGACTTTGACCGT	1094
Db	1423	GATGCCCGCTTCATTTAATTTAAAAACCAAAATTCGCGCAAAACAAGCTCTGATTGTTGATCGT	1482
Qy	1095	GAAAACTACAACAAGCTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAAGTAGGA	1154
Db	1483	GAAAAATTACAGAAGCTTTAGCTAAATTAGCTGGCGGGTTGCTGCTTTAAAGTCGGT	1542
Qy	1155	GCTCCAAACAGACAGCTTTAAAAGAAATGAAACTTTCGATTTAGGATGCTCTAAATGCT	1214
Db	1543	GCTGCACTGAACAGAAATTAAGAAATTAATAATTAACAAATTGAAGATGCATTAACGCCA	1602
Qy	1215	ACAGTCGACCGCTTGAAGAAGTATCGTGTGCTGGTGGTGAACAGACATTTATTCGGTT	1274
Db	1603	ACAGTCGCCCTCTAGAAGAAGCGATGTTTCTGCTGGTGGTACCGCACTTGCTCAATGTA	1662

RESULT 12

RESOLUT 12  
US-60-046-653-1063

: Sequence 1063, Application US/60046653

; GENERAL INFORMATION:

APPLICANT: Laqace, Robert E.

APPLICANT: Corley, Neil C.

; APPLICANT: RUSSO, Frank D.

; APPLICANT: Hann, Amy L.

; APPLICANT: Heath, Joe D.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS

; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMAC

STREET: 3174 PORT  
CITY: PLYMOUTH

CITY: PALO ALTO  
STATE: CALIFORNIA

STATE: CALIF. COUNTRY: U.S.A.

COUNTRY: U.S.  
ZIP: 94304

; ZIP: 94304  
; COMPUTER READABLE FORM:

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word perfect 6.1 for Windows/MS-DOS 6.2

; SOFTWARE: WOLF FETTER  
: CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/60/046,653

; AFFILIATION NUMBER: 03  
 ; FILING DATE: HEREWITH

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: CERRONE, MICHAEL C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PM-0003-1 P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 855-0555

; TELEFAX: (415) 845-416

; INFORMATION FOR SEQ ID NO: 1063:

; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 8784 base pairs
```

```

;
TYPE: nucleic acid

```

; STRANDEDNESS: single

TOPOLOGY: Linear

; MOLECULE TYPE: CDNA  
; IMMEDIATE SOURCE:

```

; IMMEDIATE SOURCE:
; CLONE: FEAT1G10

```

CLONE: EFALC1063  
; IIS-60-046-653-1063

05-60-040-653-1063

Query Match 56.9%; Score 945.2; DB 29; Length 8784;  
 Best Local Similarity 74.3%; Pred. No. 4.6e-244;  
 Matches 1193; Conservative 0; Mismatches 413; Indels 0; Gaps 0;

QY 15 ATGCGAAAGAAATCAAAATTTTCAGCAGATGCGCGTCTGCCATGTCGCGGAGTTGAT 74  
 DB 403 ATGCGAAAGAGATTAATTTTGCAGAGATGCGCGTCTGCCATGTCGCGGAGTTGAT 462  
 QY 75 ATGTTAGCAGATCCGCTCAAGTAAAGCTTGGTCTTAAAGGCGCAATGTTGTTCTTGA 134  
 DB 463 GTATTAGCAGATCAGTGAAGTGAATTTAGGCTTAAAGGCTGTAACGTTGTTTGA 522  
 QY 135 AAAGCTTTTGGTCTCCCTTAATTAATTAAGAGGATGATGATGATGATGATGATGAT 194  
 DB 523 AAATCATTTTGGTCTCCCTTAATTAATTAAGAGGATGATGATGATGATGATGATGAT 582  
 QY 195 TTAGAAGATCATTTTGAACAACTGGGAGCAAAATTTGGTCTCAAGTGGCTCTTAANACC 254  
 DB 583 TTAGAAGATCATTTTGAACAACTGGGAGCAAAATTTGGTCTCAAGTGGCTCTTAANACC 642  
 QY 255 AATGATATTGCTGGTATGGGAGCTACTGCAACAGCTTTTGGACAAAGCCATTTGTTCA 314  
 DB 643 AATGATATTGCTGGTATGGGAGCTACTGCAACAGCTTTTGGACAAAGCCATTTGTTCA 702  
 QY 315 GAAGGACTAAAAATGTGACAGCAGTGTGATTAATTAATTAAGGATGATGATGATGAT 374  
 DB 703 GAAGGCTTAAAAACGTAATGCTGAGGCTAACCCATTAAGTATGCTGCGGTGGGATGAA 762  
 QY 375 ACAGCAACAGCAACAGCTGTTGAAGCTTTGAAGCCATTCCTCAACCTGTATCTGGCAAG 434  
 DB 763 TTAGCAACAAACAGCAGTAGAAGAAATTAACAATATTTCATCTGTAGTTGATTCACAA 822  
 QY 435 GAAGCTATTGCTCAGGTGCGTGCAGTATCATCAGCTCTGAAAGCTTTGGAGAGTATATC 494  
 DB 823 GAAGCTATTGCTCAGGTGCGTGCAGTATCATCAGCTCTGAAAGCTTTGGAGAGTATATC 882  
 QY 495 TCAGAAGCTATTGAGGCGTGTGGCAACGATGATGATGATGATGATGATGATGATGATG 554  
 DB 883 GCGGATGCAATGGAAGTTGGTAACGAGCGGCTAATTAACCATTTGAAGAAATCAAAAGG 942  
 QY 555 ATGGAACAGAACTTCAAGTGGTTGAAGCATGCAATTTGACCGTGTACCTGCTCAA 614  
 DB 943 ATTGAACAGAAATTAGATGGTTGAAGGAATGCAATTCGACCGGTTATTATCTCAA 1002  
 QY 615 TACATGCTCAGCAGCAATGAAATATGGTTGCGAGCTTTGAAACCCATTTATCTTAATC 674  
 DB 1003 TACATGCTCAGCAGCAATGAAATATGGTTGCGAGCTTTGAAACCCATTTATCTTAAT 1062  
 QY 675 ACGGATAAAAGTGTCAACATCCAGAGATTTGCCACTACTTGGAGGATTTCTTAA 734  
 DB 1063 ACCGACAAAATATCTCAAATATTCAGATATCTTACCTTTATTAGAACAAATTTCTACA 1122  
 QY 735 ACCAACCGTCCATTTACTATTATTCAGATGATGCGGATGATGCGGATGATGCGGATG 794  
 DB 1123 CAAGCGTCCACTATTGATTTGCGGATGATGCGGATGATGCGGATGATGCGGATGAT 1182  
 QY 795 GTCTTGAACAAGATTCGTGCTACTTTCAATGTTGTTGCTGTCAAGCGCCAGGATTTGGT 854  
 DB 1183 GTATTGAACAAGATTCGTGCTACTTTCAATGTTGTTGCTGTCAAGCGCCAGGATTTGGT 1242  
 QY 855 GATCGTGTAAAGCTATGCTTGAAGCAATGCTATCTTGAAGGTTTCAGGAGTATTA 914  
 DB 1243 GACCGCGCAAGCGATGCTTGAAGATATGCTATTTTAAACAGGTTGATGATTAATCA 1302  
 QY 915 GAGGATCTAGGACTTCAATTTAAAGATGCTACAATGACAGCCCTTGGACAGCTGCTAAG 974  
 DB 1303 GACGACTTAGGTTAGAGTTTAAAGACACAACTATTGAAACCTTAGGAAATGCTAGCAA 1362  
 QY 975 ATTACAGTTGATAAGATGACAGCTAATTTGTTGAAGTTTCAGGAGTTCAGAGACTAT 1034  
 DB 1363 GTAGTTGTCGACAAAGATTAACACAACTTTGCTGAGGTTGCTGTTTCAAAAGACCAT 1422  
 QY 1035 GCTAACCGTATTGCACTGATTAATTCGCAATTTAGAAACAACTTCTGACTTTGACCGT 1094

RESULT 13

US-09-070-927-42  
 ; Sequence 42, Application US/09070927  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; APPLICANT: Patrick J. Dillon  
 ; APPLICANT: Steven C. Barash  
 ; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and  
 ; TITLE OF INVENTION: Polypeptides  
 ; NUMBER OF SEQUENCES: 982  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/070,927  
 ; FILING DATE: herewith  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: A. Anders Brookes  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3625 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 us-09-070-927-42

Query Match 56.7%; Score 942.6; DB 14; Length 3625;

Best Local Similarity 74.2%; Pred. No. 1.6e-243;  
 Matches 1191; Conservative 0; Mismatches 415; Indels 0; Gaps 0;

QY 15 ATGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTCTGCTGCGATGTCGCGGAGTTGAT 74  
 DB 364 ATGCAAAAGAGNTTAAATTTTCAGAGATGCGCGTCTGCGATGTCGCGGAGTAGAT 423  
 QY 75 ATGTTAGCAGATCCGTCAAAGTAACGCTTGGTCTTAAAGGGCGCAATGTTGTTCTTGA 134  
 DB 424 GTATTAGCAGATACAGTGAAGTGACATTAGGCGCTTAAAGGTCGTAACGTTGTTTGA 483  
 QY 135 AAAGCTTTTGGTCTCCCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194  
 DB 484 AAATCATTTTGGTTCACCATTTGATTAACGATGAGTAACAATTTGTAAGGAATTTGA 543  
 QY 195 TTAGAAGATCATTTTGAACATGCGGAGCAAAATTTGCTGAAGTGGCTTCTTAAAC 254  
 DB 544 TTGGAAGATCATTTTGAACATGCGGAGCAAAATTTGCTGAAGTGGCTTCTTAAAC 603  
 QY 255 AATGATATTGCTGGTATGGGACGACTTACTGCAACAGTTTGTGACACAGCCATTTGTCAT 314  
 DB 604 AATGATATTGCTGGTATGGGACGACTTACTGCAACAGTTTGTGACACAGCCATTTGTC 663  
 QY 315 GAAGGACTAAAAATGTCAGCAGAGTGTCTAATCCAAATTTGGTATCCGTCGAGGCAATG 374  
 DB 664 GAAGGCTTAAAAACGTTAACTGTGGAGCCAAACCCATTAGTATTCGTCGGGATTTGA 723  
 QY 375 ACAGCAACAGCAAGCTGTTGAAGCTTGAAGCCATTGCTCAACCTGTATCTGCGCAAG 434  
 DB 724 TTAGCAACAAACAGCAGTAGAAGATTACACATATTTCACTGTAGTTGATTCAAA 783  
 QY 435 GAAGCTATTGCTCAGGTCGTCAGTATCATCAGCTCTGAAAGAAATTTGGAGATATATC 494  
 DB 784 GAAGCGATTGCACAAAGTCGCTGTTTCATCAGGTTCTGAAAGAAATTCGCGCAATTA 843  
 QY 495 TCAGAACTATGAGCGTGTGGCAACGATGCTGATTAACCATCGAAGATCTCGAGGT 554  
 DB 844 GCCGATGCAATGGAAGTTGGTAAAGCGGCGTAATTAACCATTTGAAGAAATCAAAAG 903  
 QY 555 ATGGAACAGAACTTGAAGTGGTTGAAGGATGCAATTTTGACCGTGGTTACCTGCTCAA 614  
 DB 904 ATTGAACAGAAATTAGATGTTGTTGAAGAAATGCAATTCACCGCGTTATTTATCTCAA 963  
 QY 615 TACATGCTCAGACAAATGAAAAATGGTTGACAGCTTGAAGAAATTCATTTATCTTAATC 674  
 DB 964 TACATGCTTACTGACACAGTAAATGGAAGCTGTTTGAAGAAATTCATTTATCTTAAT 1023  
 QY 675 ACGGATAAAAAAGTGTCAACATCCAGACATTTTGGCACTACTTGAAGAGTTCTTAA 734  
 DB 1024 ACCGACAAAAAATCTCAAAATTTCAAGATATCTTACCTTTATTAAGAAATTTCTACA 1083  
 QY 735 ACCAACCGTCCATTTACTTATTATTCAGATGATGATGATGATGATGATGATGATGATG 794  
 DB 1084 CAAGCCGCTCCACTTATTTATTCGAGATGATGATGATGATGATGATGATGATGATG 1143  
 QY 795 GTCTTGAACAAAGATTCGTGGTACTTTCATATGTTGCTGTCAAGGCGCAGGATTTGGT 854  
 DB 1144 GTATTGAACAAATCCGTGCTACATTTAATGTTGTCGACGATAAAGCGCAGGATTTGG 1203  
 QY 855 GATCGTCGTAAGATATGCTTTGAAGACATTTGCTATCTTTGACAGGTGCTACAGTATCA 914

DB 1204 CACGCCGCAAGAGGATGCTTGAAGATATTGCTATTATAACAGGTGTACAGTAATCACT 1263  
 QY 915 GAGGATCTAGGACTTGAATTAAGAGATGCTACAAATGACAGCCCTTGGACGGCTGCTAAG 974  
 DB 1264 GACGACTTAGGTTAGAGTTAAAGACACAACTATTGAAATCTTAGGAAATGCTAGCAA 1323  
 QY 975 ATTACAGTTGATAAAGATAGCAGATTAATTTGTTGAAGTTTCAGGAAGTTTCAGAGCTAT 1034  
 DB 1324 GTAGTTGTCGCAAAAGATAACACAACTTTGTCGAAGTGTCTGTTCAAAAGAGCCATT 1383  
 QY 1035 GCTAACCGTATTGCACTGATTAATTCGAATTAAGAAACAACTTCTGACTTTGACCGT 1094  
 DB 1384 GATCCCGCGTTTCAATTTAAATTAAGAAACAACTCGGAAACAACTGCTGATTTGATCGT 1443  
 QY 1095 GAAAACTACAAGAACTTTGGCGAAATTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1154  
 DB 1444 GAAAAATTAACAAGAACGTTTAGCTTAATTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1503  
 QY 1155 GCTCCAAACAGACAGCTTTTAAAGAAATGAACCTTCGCAATTTAGAGTGTCTTAAATGCT 1214  
 DB 1504 GCTCAACTGAACAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1563  
 QY 1215 ACAGTCCAGCCGTTGAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274  
 DB 1564 ACAGTCCCGCTGTAGAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1623  
 QY 1275 ATTAAAAAGTAGCAGCTTTGAGCTTTGAGGCGCATGATGCTGCTGCTGCTGCTGCTGCTG 1334  
 DB 1624 ATTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683  
 QY 1335 CTTGCTGCTCTAGAGAGCCCTGTACGTTCAATTTAAATGCTGCTGCTGCTGCTGCTGCTGCT 1394  
 DB 1684 GTTCGTCATTTAGAGAACCAATCCGTTCAATTCGTTGAAATGCTGCTGCTGCTGCTGCTG 1743  
 QY 1395 GTAGTTATTGACAAAGTTGAAAAACAGCCCTGCGAGGACAGGATTTAATGCTGCAACAG 1454  
 DB 1744 GTGATTTGTCACAACTTAAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1803  
 QY 1455 GAGTGGTGTGATGATTTAAACAGGAATCATTTGACCTGTCAAAGTAACACGATCAGCG 1514  
 DB 1804 GAATGGTAAACATGTTGAAGCCGTTATGTTGACCCCAACAAAGTAACGCTGCTGCTG 1863  
 QY 1515 CTTCAAAATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1574  
 DB 1864 TTACAAATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923  
 QY 1575 AAACCTGAACAGCTACGCCAGCCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
 DB 1924 AAACCAAGAACAGCTGACCCAGCTCTTATGATGATGATGATGATGATGATGATGATG 1969

## RESULT 14

US-09-070-927A-42

; Sequence 42, Application US/09070927A

; GENERAL INFORMATION:

; APPLICANT: Charles A. Kunsch

; Patrick J. Dillon

; Steven Barash

; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

; NUMBER OF SEQUENCES: 982

; CORRESPONDENCE ADDRESS:

; ADDRESS: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/070, 927A  
 ; FILING DATE: 04-May-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/046,655  
 ; FILING DATE: 1997-05-16  
 ; APPLICATION NUMBER: 60/044,031  
 ; FILING DATE: 1997-05-06  
 ; APPLICATION NUMBER: 60/066,009  
 ; FILING DATE: 1997-11-14  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kenley K. Hoover  
 ; REGISTRATION NUMBER: 40,302  
 ; REFERENCE/DOCKET NUMBER: PB369  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 42:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3625 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 42:  
 ; US-09-070-927A-42

Query Match 56.7%; Score 942.6; DB 14; Length 3625;  
 Best Local Similarity 74.2%; Pred. No. 1.6e-243;  
 Matches 1191; Conservative 0; Mismatches 415; Indels 0; Gaps 0;  
 QY 15 ATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTCTGCCATGCTGCGCGGAGTTGAT 74  
 DB 364 ATGGCAAAAGAGTTAAATTCAGAGAGATGACGTCGACGATGTCGCGGAGTAGAT 423  
 QY 75 ATGTTAGAGATACCTCAAGTAAGCTTGGTCTTAAAGGGGCAATGTTGTTCTTAA 134  
 DB 424 GTATTAGCAGATACAGTAAAGTGACATTAGGCGCTTAAAGGTCGTAAACGTTGTTAGAA 483  
 QY 135 AAAGCTTTTGGTCTCCCTTAATTAATGATGAGGGGTAAACCATTCCTAAAGAGATCGAA 194  
 DB 484 AATCAATTTGGTTCACCTTATTAATGATGAGGGGTAAACCATTCCTAAAGAGATCGAA 543  
 QY 195 TTAGAAGATCATTTTGAACATCGGAGCAAAATTTGCTGCTCAAGTGGCTCTTAAACCC 254  
 DB 544 TTGGAAGATCATTTTGAACATCGGAGCAAAATTTGCTGCTCAAGTGGCTCTTAAACCC 603  
 QY 255 AATGATATTGCTGGTATGGAGGACTACTGCAACAGTTTTCAGCAAGCCCATTTGTTTAT 314  
 DB 604 AATGATATTGCTGGTATGGAGGACTACTGCAACAGTTTTCAGCAAGCCCATTTGTTTAT 663  
 QY 315 GAAGGACTAAAGATGACAGAGTGTAAATCCAAATTTGGTATCCGTCGAGGCAATTTGAA 374  
 DB 664 GAAGGCTTAAAGATGACAGAGTGTAAATCCAAATTTGGTATCCGTCGAGGCAATTTGAA 723  
 QY 375 ACAGCAACAGCAGCAGCTTTGAAGCCTTTGAAGCCTTTGCTCAAGCTGCTATCTGGCAAG 434  
 DB 724 TTAGCAACAAACAGCAGCTAGAGAAATTTACAAATTTTATCTCTAGTTGATTTCAAA 783  
 QY 435 GAAGCTATTGCTCAGTCTGCTGATATCATCATCGCTCTGAAAGAGTTGGAGAGATATATC 494  
 DB 784 GAAGCATTTGCAACAGTCTGCTGTTTATCAGGTTCTGAAAGAGTCTGGCCCAATTAAT 843  
 QY 495 TCAGAGCTATGAGGCTGTGGCAGCAGTGTGATTAACCATCGAAATTCGAGT 554  
 DB 844 GCGATGCAATGGAAGAGTTGGTACGAGCGGTAAATTTACCATTTGAAGAAATCAAAAGG 903  
 QY 555 ATGGAACAGAGATTTGAAGTGTGAAGCAGTCAATTTGACCGTGTACCTGCTCAAA 614  
 DB 904 ATTGAACAGAGATTTGATGCTGTGAAGAGATGCAATTTGACCGGTTATTTATCTCAA 963  
 QY 615 TACATGGTCACAGACANTGAAAAATGGTTGACAGCTTTGAAACCCATTTATCTTAATC 674

DB 964 TACATGGTTACTGACAAACGATAAAATGGAAGCTGTTTGAAGAAATCCATATCTTAAT 1023  
 QY 675 ACCGATAAAAAAGTGTCAACATCCCAAGACATTTTCCACTACTTCCAGGAAGTCTCTAAA 734  
 DB 1024 ACCGATAAAAAATCTCAATATTTCAAGATATCTTACCCTTTATTGAACAAATTTACAA 1083  
 QY 735 ACCAACCGTCCATTACTTACTTATTTGAGATGATGTTGAGTGTGAAGCAGCTTCCAAACCTT 794  
 DB 1084 CAAAGCGTCCACTATTGATTTGCGGATGATTTGATGGGAAAGCTCTACCAACATTA 1143  
 QY 795 GTCTTGAACAAGATTCGTGGTACTTCAATGTTGTTCTCAAGCGCCAGGATTTGGT 854  
 DB 1144 GTATTGAACAAATTCGTTGATCAATTTGTCGAGTAAAGCGCCAGGATTTGGT 1203  
 QY 855 GATCGTCTGTAAGCTATGCTTGAAGACATTTGCTATCTTACAGAGTGTGACAGTATTACA 914  
 DB 1204 GACCGCGCAAGCGATGCTTGAAGATTTGCTATTTTAAACAGGTGTGACAGTAACTACT 1263  
 QY 915 GAGGATCTAGGACTTGAATTTAAAGATGCTACAATGACAGCCCTTGGACAGGCTGCTAAG 974  
 DB 1264 GACGACTTAGGTTAGAGTTAAAGACACAACTATTGAAACTTTAGAAATGCTACCAA 1323  
 QY 975 ATTACAGTTGATAAGATAGACAGATAATTTGTTGAAGTTTCAGGAAGTTTCAGAACTATT 1034  
 DB 1324 GTAGTTGTCGACAAAGATAACACAACTATTCGAAGGTGCTGGTTCAAAAGAGCAATT 1383  
 QY 1035 GCTAACCGTATTGCACTGATTAATTCGCAATTTAGAAACAACTTCTGACCTTTGACCGT 1094  
 DB 1384 GATGCGCGGCTTCAATTTAAATTAACCAACGCGGCAACACGCTGCTGATTTGATCT 1443  
 QY 1095 GAAAACTACAAGAGCTTTGGCGAAATTTAGCTGGTGGTGTAGCTGTTTATCAAGATGGA 1154  
 DB 1444 GAAAAATTTACAGAGCTTTAGCTAAATTTAGCTGGTGGTGGTGGTGGTGGTGGTGGT 1503  
 QY 1155 GCTCCAAAGAGACAGCTTTTAAAGAAATGAACTTTGCAATTTGAGGATGCTCTAAATGCT 1214  
 DB 1504 GCTGCAACTGAAACAGAAATTTAAAGAAATTTAAATTTACGAATTTGAAGTGCATTTAAACGCA 1563  
 QY 1215 ACAGCTGACGCGTTCAGAAAGTATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1274  
 DB 1564 ACAGCTGACGCGTTCAGAAAGTATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1623  
 QY 1275 ATTGAAAAAGTAGCAGCTCTTGAAGCTTGAAGGCGATGATGCTACTGGAGCTTAACATTGTG 1334  
 DB 1624 ATTGTTAAAGTCGCTGCGCTAGAGCTGAAGCGGATGTCGCAACAGGATCAAGATTGTC 1683  
 QY 1335 CTTGCTGCTCTAGAGAGCTCTGATGCTCAATTTGCTTTAAATGCTGGGTACGAAGGCTCC 1394  
 DB 1684 GTTCTGCTATTAGAGAACCAATTCGCTCAATTCGCTGAAATGCTGGTTATGAAGGATCA 1743  
 QY 1395 GTAGTTATTGCAAGTTTGAAGAACCCCTTGCAGGACAGGATTTAATGCTGCAACAGT 1454  
 DB 1744 GTAGTTATTGCAAGTTTGAAGAACCCCTTGCAGGACAGGATTTAATGCTGCAACAGT 1803  
 QY 1455 GAGTGGTGTGATGATTAATTAACAGGAATTCATTGACCCCTTGCAGGACAGGATTAACAGT 1514  
 DB 1804 GAATGGTAAACATGTTGCAAGCCGCTATTGTTGACCCCAACAAAGTAACTGCTTCTGCC 1863  
 QY 1515 CTTCAAAATGACGCTCTGATGCTAGTCTTATTTTGAACAAAGAGGATGCTTGTCTAAT 1574  
 DB 1864 TTACAAATGACGCTCTGATGCTAGTCTTATTTTAAACAACTGGAAGCAGTGTGCTGAGAC 1923  
 QY 1575 AAACCTGAACAGCTTACGCCAGCGCCAGCAATGCGCAGGAGTATGG 1620  
 DB 1924 AAACAGAACAGCTGACCCAGCTCTCTATGATGGATCCATCAATGG 1969

RESULT 15  
 US-09-134-000-1154  
 ; Sequence 1154, Application US/09134000A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS  
; FILE REFERENCE: GTC-005  
; CURRENT APPLICATION NUMBER: US/09/134,000A  
; CURRENT FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 6810  
; SEQ ID NO 1154  
; LENGTH: 1638  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis  
US-09-134-000-1154

Query Match 56.2%; Score 934; DB 15; Length 1638;  
Best Local Similarity 73.8%; Pred. No. 2.4e-241;  
Matches 1186; Conservative 0; Mismatches 420; Indels 0; Gaps 0;  
QY 15 ATGCGAAACAAATCAAAATTTTCAGCAGATGCGCGTGCTGCCATGTCGCGGAGTGTGAT 74  
DB 13 atgcaaaagagattaatttcgagaagatgcacgfcgacgaatgctcacgcgagtagat 72  
QY 75 ATGTTAGCAGATACCGCTCAAGTAACGCTTGGTCCTAAAGGGGCGCAATGTTGTTTGA 134  
DB 73 gtattagcagatcacgtgaaagtgacattagccctaaagctgtaacgtttgttttagaa 132  
QY 135 AAGCTTTTGGTCTCCCTTAATTACTAATTAATGACGGGTAAACCAATGCTAAAGAGATCGAA 194  
DB 133 aaatcattggttcaccattgattactaacgatggagtaacaattgctaaggaaatgaa 192  
QY 195 TTAGAAGATCATTTTGAACATGCGGAGCAAAATTTGGTCTGCTGAAGTGGCTCTTAAACC 254  
DB 193 tggagagatcatttgaacaatg99gcgaatattagttccagaagtgtctctcaaacg 252  
QY 255 AATGATATTGCTGATGGAGGAGTACTGCAACAGTCTTGGTACACAAAGCCATTGTTTCAT 314  
DB 253 aatgatatgctggtgacggaacaacaacagcagctgttttgacaaagccattgttcgt 312  
QY 315 GAAGGACTAAAAATGACAGCAGGTGCTTAATCCAAATTTGGTATCCCTGACGAGCATTTGAA 374  
DB 313 gaaggcttaaaaaacglaactgctggagccaaacccattaggtattcgtcggtggtgaa 372  
QY 375 ACAGCAACACACAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAG 434  
DB 373 ttgacaaaacacagcagtagaagaattacacaatttctcattctgtattgttcaaaa 432  
QY 435 GAAGCTATTGCTCAGGTCGCTGAGTATCATCATCGCTCTGAAAAAGTTGGAGATATATC 494  
DB 433 gaagcattgcacaagtcgctgctgttctcaggttctgaaaaagtcggccaaattaatt 492  
QY 495 TCAGAGCTATGAGCGTGTGGGCAACGATGGTGTGATTACCATTCGGAAGATCTCGAGGT 554  
DB 493 gccgaTgcaatggaaaaagtgtgtaacgcgcgcgtaattaccattgaagaatcaaaaagg 552  
QY 555 ATGGAACACAGCTTGAAGTGGTGAAGGCATGCAATTTGACCGTGGTTACCTGTCTCAA 614  
DB 553 attgaaacagaaatagattg99gtggaaggaatgcaattcgaacgcgcggtatttctcaa 612  
QY 615 TACATGCTCAGACAAATGAAAAATGGTTGCGAGCCTTGAACCCATTTATCTTAAATC 674  
DB 613 tacatggttactgacaacgataaaatggaagctgttttagaaaatccatatcttaatt 672  
QY 675 ACCGATAAAAAAGTGTCAACATCCCAAGACATTTTGGCACTACTGTAGGAGTTCCTTAAA 734  
DB 673 accgacaaaaaaactcaaatattcaagatatcttaacctttattagaaacaattctcaaa 732  
QY 735 ACCAACCGTCCATTACTTATTATTCAGATGATGTGATGGTGAAGCATTCCAAACCCCTT 794  
DB 733 caaagcgcctcattgattgtg99gtggaagtgatgttggggaagcgtctaccaacatta 792  
QY 795 GTCTTGAACAAGATTCGTGGTACTTTCAATGTGGTCTGCTCAAGGCGCCAGGATTTGGT 854  
DB 793 gtattgaacaaaatccgtggttacatttaattg99gtgcgcagataaaaagcgcgaggtttggt 852

QY 855 GATCGCTGTAAGCTATGCTTGAAGACATTTGCTATCTTTGACAGTGGTGTACAGTGAATACA 914  
DB 853 gaccgcgcaaacgagatgcttgaagatattgctattttaacagtggtgacagtaact 912  
QY 915 GAGGATCTAGGACTTGAATTAAGATGCTACATGACAGCCCTTGGACAGCTGCTAAG 974  
DB 913 gacgacttagg99gttag99gttaaaagcacacaactattgaaaacttag99aaatgctcagcaaa 972  
QY 975 ATTACAGTTGATAAAGATAGCACAGTAATTTGTAAGGTTTCAGGAAGTTTCAGAGCTATT 1034  
DB 973 gtagttgctgacaaaagataacacaacaattgctgaaggtgctggttcaaaaagaagccatt 1032  
QY 1035 GCTAACCTATTGCACTGATTAAATCGCAATTAGAAACACAACTTCTGACTTTGACCGGT 1094  
DB 1033 gatgccgcgcttcatttaattaaacacaaatcgcgcaaacacgctgattgattgactcgt 1092  
QY 1095 GAAAACTACAAAGAACGTTTGGCGAAAATTAGCTGCTGGTGTAGCTGTTATCAAAAGTAGGA 1154  
DB 1093 gaaaaattacaagaacgttttagctaaattagctg99gtg99gtcgtttaaagtcggt 1152  
QY 1155 GCTCCAAACAGAGACAGCTTTAAAGAAATGAAACTTCGCATTGAGGATGCTCTAAATGCT 1214  
DB 1153 gctgcaactgaacagaaattaaagaattaaataattacgaattgaagatgattaaacgca 1212  
QY 1215 ACAGCTCAGCCGTTGGAAGAGGTATCGTTCTGCTGGTGGGAACAGCACACTTATTACGGTT 1274  
DB 1213 acacgtgccctgtaaaaaagcagctgttctg99gtg99gtacccgacgtgcaatgta 1272  
QY 1275 ATTCAAAAGTAGCAGCTCTTGAGCTTGAGGGCGATGATGCTACTGACGCTTAACATTGTG 1334  
DB 1273 attg99taaagtcgctgc99tagaagctgaag99gagatgtgcaacacagggatcaaaattgct 1332  
QY 1335 CTTGCTGCTCTAGAAGAGCCTGTACGTCAAATTTGCTTTAAATGCTGGGTACGAAGGCTCC 1394  
DB 1333 gttcgtgcattaaagaagaccaaaccgtaaatcgtaaatgctg99gttatgaaggatca 1392  
QY 1395 GTAGTTATTGACAAAGTTGAAAAACAGCCCTGCGAGAACAGGATTTAATGCTGCAACAGGT 1454  
DB 1393 gtgattgtgacaaactaaaaaattgtgacttaggtatcgatatacaatgcaatgcaatgcaatg 1452  
QY 1455 GAGTGGCTTGTATGATTAAACAGGAATCATTTGACCTGTCAAAAGTAACACGATCAGCG 1514  
DB 1453 gaatgg99taaacaatg99gtgaaacccggtattgttgaacccatcaaaagtaactcgtctgcc 1512  
QY 1515 CTTCAAAATGCACTTCTGTAGCTAGTCTTATTGACACAGAGAGGATTTGCTGCTAAAT 1574  
DB 1513 ttacaaaatgcaagcttctgctcagctttattattaaacaaactgaagcaattgttgcagac 1572  
QY 1575 AAACCTGAACAGCTACGCGCAGCCGCAATGCCAGCAGTATGG 1620  
DB 1573 aaaccagaaccagttgaccacgctcctatgatggatccatcaatgg 1618

Search completed: February 15, 2001, 22:42:58  
Job time: 7781 sec

**THIS PAGE BLANK (USPTO)**